EFFECTIVE INTERVENTION THROUGH SELF-MANAGEMENT: A META-ANALYSIS

A Dissertation

by

JOHN LOCKETT DAVIS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee	Kimberly J. Vannest
Co-Chair of Committee	William A. Rae
Committee Members	Cecil R. Reynolds
	Richard I. Parker
	Dominique T. Chlup
Head of Department	Victor Willson

August 2015

Major Subject: School Psychology

ABSTRACT

The current study examines aspects of Self-managed (SM) interventions for behavior. SM is a multi-component intervention where students are active participants in the process of improving their behavior. Despite substantial research supporting the use of this intervention in schools, there is currently a lack of consensus regarding which procedures are necessary to implement this intervention. Methods for evaluating *evidence based practice* (EBP) place an emphasis on clearly identifiable intervention procedures. Therefore, a review of SM literature is necessary to determine if sufficient supportive evidence exists to promote various versions of SM as an EBP. In addition, an analysis is necessary to determine which intervention components are essential and how the addition or removal of components relates to student outcomes.

The current study is a meta-analytic review of single case research literature. In reviewing the presence of intervention components, the current study found that 18 unique combinations of intervention components occur within SM literature. When disaggregated based on student responsibility for implementing intervention components, the number of unique combinations of intervention procedures grew to 44. Application of current EBP criteria found that four of versions of the SM intervention met criteria.

Examination of student factors found differential effects based on student disability classification, age, and the targeted outcome. No effects were found within the educational setting variable.

Examination of additional intervention procedures found that the use of external reinforcement and assessment of student accuracy is not related to improved outcomes. In addition, the method of cueing self-recording is not related to differential effects.

In sum, the current analysis shows that SM is a highly effective intervention for school age children across a variety of behavioral targets. In terms of implementation, the current analysis shows that much of SM's effects are based two key components (e.g. self-assessment and self-recording). The current analysis also shows that many of the procedures commonly associated with this intervention do not contribute to improved outcomes.

DEDICATION

I would like to dedicate this work to my wife Heather and twin daughters, Ruby and Scarlett. Your love, support, and sacrifice has helped me more that you will ever know.

ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs Dr. Vannest and Dr. Rae for their support in completing this research. I would also like to thank my committee members, Dr. Parker, Dr. Reynolds, and Dr. Chlup for their time and insight throughout this process.

I would like to give a special acknowledgement to Dr. Vannest and Dr. Parker for their mentorship throughout my doctoral program. I am truly grateful for the opportunity you both afforded me in learning the academic process.

I would like to thank Dr. James McNamara for his extremely generous contribution to my education. The two and a half years we spent one-on-one studying non-parametric statistics was truly a gift for me. Our time together helped me to learn a more elegant and flexible orientation to the analysis of data.

Finally, I would like to thank my parents Drs. John and Barbara Davis for your unique and substantive contributions to this project. I am very fortunate to two great scholars as parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

School personnel are often the first to respond to the emotional and behavioral needs of school-age children (Burns et al., 1996). Given this responsibility, school-based practitioners should be prepared to address a variety of emotional and behavioral difficulties (Harrison, Vannest, Davis, & Reynolds, 2012). However, intervening with behaviors that interfere with learning is a complex task that requires specialized skills (Brouwers & Tomic, 2000). Further, current accountability standards for teachers place the highest importance on academic skill attainment rather than remediation of problem behavior (Vannest, Mahadevan, Mason, & Temple-Harvey, 2009). As a result, students' social, emotional, and behavioral needs are often neglected (Kataoka, Rowan, & Hoagwood, 2009; Landrum, Tankersley, & Kaufman, 2003).

Although a variety of empirically validated interventions are available for school-age children (Vannest, Reynolds, & Kamphaus, 2008), a teacher's ability to intervene is often compromised by the diverse and challenging responsibilities inherent in managing a classroom (Franz et al., 2008).

Given the demands of the classroom, teachers are in need of effective and empirically sound intervention tools that can be integrated into the classroom routine. Self-management (SM) interventions have the potential to fulfill this need (Hughes, Ruhl, & Misra, 1989; Lee, Simpson, & Shogren, 2007; Reid, 1996). SM is a multicomponent intervention that relies on a variety of student- and teacher-directed activities (Mace, Belfiore, & Hutchinson, 2001).

SM has been found to be an effective behavioral intervention in schools (Briesch & Chafouleas, 2009; Fantuzzo & Polite, 1990; Webber, Schuermann, McCall, & Coleman, 1993) with positive effects across settings (Hoff & DuPaul, 1998; McDougall & Brady, 1995; Smith & Sugai, 2000; Wood, Murdock, & Cronin, 2002) and behavior constellations (Barry & Messer, 2003; Rooney & Hallahan, 1988; Strain, Kohler, Storey, & Danko, 1994).

Despite the promising evidence promoting SM as an intervention for behavior, there is no clear consensus on which specific intervention procedures are necessary for effective implementation of this intervention. Thus, different combinations of intervention components have been examined (Briesch & Chafouleas, 2009; Fantuzzo & Polite, 1990; Fantuzzo, Rohrbeck, & Azar, 1987). For example, Fantuzzo and colleagues (1987) created a conceptual framework for the components of SM interventions and identified 11 unique components to be considered when implementing SM. These may include: “selection and definition of target behavior, goal setting, observation or recording of target behavior, evaluation of goal attainment, selection of reinforcers, instructional prompts, administration of secondary and primary reinforcers, and graphing or charting behavior” (Fantuzzo & Polite, 1990, p. 182). Given the breadth of potential components, applications of the Fantuzzo et al. (1987) framework have shown considerable variation among research studies that apply SM interventions. (Fantuzzo & Polite 1990; Briesch & Chafouleas, 2009). This lack of consistency in the literature raises questions regarding the structure of this intervention, what components are necessary for implementation, and how student outcomes are affected by different combinations of components.

Two meta-analyses have examined differences in student outcomes based on the number of components used in SM interventions (Briesch & Chafouleas, 2009; Fantuzzo et al., 1987). Neither of these analyses found differences in intervention effects based on the total number of intervention components. That is, SM interventions with lower total components showed similar effects to interventions with higher component totals. Thus, these findings support the potential benefits of a more parsimonious SM intervention.

While the examination of component totals is important in understanding this intervention, more research is necessary to examine the relative contribution of individual components. Evaluation of total numbers of components may mask useful information regarding the utility of certain components, as two intervention packages with the same total may be composed of two entirely different sets of components. In addition, a more specific evaluation of intervention components will inform interventionists regarding which components are necessary for implementing this intervention and which components may provide additive effects. In short, a more uniform approach to this intervention is needed to permit replication and establish a clear evidence base (Baer, Wolf, & Risley, 1968; Chambless & Ollendick, 2001).

CHAPTER II

REVIEW OF THE LITERATURE

Self-management (SM) is an intervention growing out of the self-regulation work of Bandura (1969, 1977). Generally, this intervention is viewed as a multistage process consisting of student-directed evaluation and recording of relevant behavior (Mace et al., 2001).

Several theoretical models have been posited to explain student *reactivity* or responsiveness to this intervention. Operant theorists assert that the behavior change that occurs with this intervention can be explained by the self-administered cues an individual uses to sensitize himself to environmental consequences (Mace, Shapiro, West, Campbell, & Altman 1986; Snider, 1987). According to this theory, self-administered behavioral assessment and consequences are contingent upon the activity of self-assessment and self-recording (Rachlin, 1974). Therefore, only student-directed activities act to change behavior.

In an extension of operant theory, Nelson and Hayes (1981) proposed that other agents act in connection with the student's self-directed processes to affect behavior change. Indeed, within operant theory, the entire SM process, including cues from the teacher, the recording device, and other environmental stimuli, is seen to cause behavior changes (Nelson & Hayes, 1981).

In contrast to the purely behavioral model, Kanfer (1977) proposed a three-stage model based on cognitive-behavioral theory to explain reactivity. In this model, the processes of self-monitoring, self-evaluation, and self-reinforcement act to guide behavior change. Through self-monitoring and self-evaluation, the student becomes conscious of relevant behaviors in relation to an external set of expectations. If the student's behavior matches or exceeds a predetermined criterion, self-reinforcement occurs and this reinforcement, in turn, acts to influence external expression of behavior (Kanfer, 1977). The assumption in this model is that internal cognitive processes (i.e., self-evaluation) can influence observable behaviors (Mace & Kratochwill, 1985).

The operant and cognitive-behavioral models diverge in the emphasis placed on environmental consequences in affecting behavior change (Mace et al., 1986). However, both models place primary importance on students' evaluation of their own behavior as the primary agent of change.

Application of Self-Management

SM intervention packages have a history of effective application (Briesch & Chafouleas, 2009; Fantuzzo & Polite, 1990; Hughes et al., 1989; Lee et al., 2007; Mooney, Ryan, Uhing, Reid, & Epstein, 2005; Reid, 1996; Snider, 1987; Webber et al., 1993). However, there is no consensus on the "active ingredients" that create the most effective intervention. Directing more attention to specific aspects of SM is important in fully evaluating this intervention. For purposes of application, it is important to be able to identify and describe the various components that make up a particular behavioral intervention. Early behavior interventionists emphasized the need to clearly define intervention techniques to promote replicability (Baer et al., 1968). This consideration has been adopted in more recent paradigms for evaluating evidence-based practices

(EBP) through an emphasis on clearly identified intervention processes (Chambless & Ollendick, 2001).

In recent years, specific criteria have emerged for evaluating EBP in single-case research (Horner et al., 2005, Horner & Kratochwill 2012; Kratochwill et al., 2010). These criteria call for minimum standards for study design, minimum amounts of data collected, and evidence of reliable measurement. In addition to minimum design and data standards, contemporary EBP standards call for specific information regarding the composition of the intervention, alignment among the population/populations of students the intervention is intended to support, and evidence of continuity in the positive outcomes the intervention is intended to produce. Thus, according to Horner and Kratochwill (2012), an EBP should have an "operationally defined set of procedures that are used by a specified target audience, under defined conditions/contexts, to achieve valued outcomes for one or more defined populations" (p. 267).

The current review will examine the extant literature base to evaluate how specific intervention components impact intervention results. This procedural review will help to evaluate the effectiveness of various SM intervention protocols. In addition, this analysis will evaluate SM procedures to determine which sets of procedures meet criteria for classification as an EBP.

Defining Self-Management

Of primary importance in promoting SM as an EBP is defining the intervention. However, creating a set of "operationally defined procedures" is difficult given the variability in implementation noted within published research. For example, in the past, the descriptive label "self-management" covered several different intervention protocols and components (Briesch & Chafouleas, 2009; Fantuzzo et al., 1987; Fantuzzo & Polite, 1990; Mace et al., 2001).

Specifically, throughout the history of SM implementation, four general types of interventions have been identified (see Table 1): self-monitoring (Mace et al., 2001); self -evaluation (Mace et al., 2001); self-instruction (Bornstein & Quevillon, 1976; Meichenbaum & Goodman, 1971); and self-reinforcement (Barling & Patz, 1980; Morris & Messer, 1978). Although specific intervention procedures are available for these four intervention protocols, applied researchers rarely adhere to these categorizations. For example, a study using self-monitoring methodology may employ a self-instruction component (Hallahan, Lloyd, Kneedler, & Marshall, 1982), a self-reinforcement component (Smith, Young, West, & Morgan, 1988), or all three components simultaneously (Ninness, Fuerst, Rutherford, & Glenn, 1991). Given this overlap, an important goal is to define and test outcomes from each intervention package to better understand the relative contribution of each intervention component.

Table 1

Self-Management Interventions

Intervention	Definition
Self-Monitoring	A multiple-step process whereby the student observes the occurrence or non-occurrence of the behavior and records features of the behavior (Mace et al., 2001).
Self-Evaluation	An intervention in which a student is involved in the determination and/or evaluation of a performance goal (Mace et al., 2001).
Self-Instruction	An intervention based on student-generated instructional statements. In this intervention, the student applies predetermined self-directed statements to guide behavior (Bornstein & Quevillon, 1976; Meichenbaum & Goodman, 1971).
Self-Reinforcement	An intervention in which the student chooses and administers external reinforcement when a predetermined criterion is met (Barling & Patz, 1978; Morris & Messer, 1978).

Number and type of components. Taking a structuralist approach to the identification of intervention components related to SM, Fantuzzo and colleagues identified 11 potential components that may be present (see Table 2; Fantuzzo et al., 1987). These may include “selection and definition of target behavior, goal setting, observation or recording of target behavior, evaluation of goal attainment, selection of reinforcers, instructional prompts, administration of secondary and primary reinforcers, and graphing or charting behavior” (Fantuzzo & Polite, 1990, p. 182).

Table 2

Potential Components and Definitions

Component	Definition
Selection of dependent variable	Selecting and prioritizing behavioral outcome variable(s)
Definition of target behavior	Creating an operational definition of target behavior
Determination of performance goal	Determining the performance criteria for target behavior
Evaluation to determine whether performance goal was met	Comparing actual performance of target with stated performance goal
Instructional prompts for target behavior	Delivering prompt(s) to engage in target behavior
Observation of target behavior	Making a judgment as to the presence or absence of target behavior
Recording	Documenting the occurrence of target behavior
Graphing or charting behavior	Summative documentation of the student recording, either graphically or with written notes
Selection of primary reinforcer	Choosing the primary reinforcer

Table 2 (Continued)

Potential Components and Definitions

Component	Definition
Administration of primary reinforcer	Dispensing or initiating the dispensation of primary external reinforcers
Administration of secondary reinforcer	Dispensing of tokens or points to be exchanged for primary reinforcement

Given the wide range of potential components employed within SM, researchers have varied considerably in their implementation of intervention components. To further confound a uniform understanding of SM, student involvement within certain components of intervention has also varied across studies (Fantuzzo et al., 1987; Fantuzzo & Polite, 1990).

In an attempt to unify the literature base on SM intervention components, Fantuzzo and colleagues (1987) conducted a review, and later a meta-analysis (Fantuzzo & Polite, 1990) comparing studies based on the total number of intervention components used. No significant difference was found in effects between studies based on the total number of components used in the intervention (Fantuzzo et al., 1987; Fantuzzo & Polite, 1990).

A more contemporary meta-analysis (Briesch & Chafouleas, 2009) updated Fantuzzo and colleagues' work, using articles published between 1988 and 2008. The authors found identical results, confirming the idea that the total number of intervention components does not affect outcomes. Further, Lee et al. (2007) examined the "number of components" variable in a meta-analysis of SM interventions for students with autism. Within this more narrow population of students, Lee and colleagues also found no difference between interventions that involved components termed self-monitoring (i.e., self-assessment and self-recording) or self-evaluation (i.e., decision making and goal-setting). Each of these meta-analyses individually support the finding that the total number of components is not related to differentiated effects for this intervention.

With evidence that the total number of components appears to have no discernible difference on outcomes (Briesch & Chafouleas, 2009; Fantuzzo et al., 1987; Fantuzzo & Polite, 1990), the next logical step in this line of inquiry is to examine the specific methodological issues that may impact intervention outcomes. Although past meta-analyses have addressed the issue of whether "more is better," it is still not clear how specific components of an SM intervention affect outcomes. Since the early 1970s, several components have been debated relative to their necessity in implementing SM (Snider, 1987; Webber et al., 1993). These components are detailed in the following sections.

Student involvement. Student involvement is posited as a major contributor to the effects of SM interventions (Hallahan et al., 1982; Snider 1987). However, the level of student-directed implementation and responsibility in SM intervention is highly variable across studies (Briesch & Chafouleas 2009; Fantuzzo et al., 1987). The theoretical underpinnings of this intervention maintain that student-initiated intervention builds self-awareness of behavior and teaches the ability to self-regulate (Hallahan et al., 1982; Snider, 1987). Thus, the expectation is that following this intervention, the student will have gained the skills to manage their behavior independently (Snider, 1987). Therefore, the responsibility for appropriate behavior would likely affect the student's ability to learn and internalize self-awareness of the targeted behavior (Snider, 1987). For example, several research groups have implemented SM interventions that relied on self-assessment and self-recording (Amato-Zech, Hoff, & Deopke, 2006; Broden, Hall, & Mitts, 1971; Christie, Hiss, & Lozanoff, 1984; Hughes et al., 2002; Lloyd, Hallahan Kosiewicz, & Kneedler, 1982), whereas other studies have utilized the same intervention procedures but gave the responsibility of self-recording to someone other than the student (e.g., teacher or researcher) (Hughes et al., 2002; Lloyd et al., 1982). Given this change, it is important to determine if these modifications in implementation responsibility resulted in an effect on the targeted outcome.

With the known variability in implementation of the 11 identified components, the addition of the student responsibility to any or all of these components adds a level of complexity to understanding and operationalizing the procedures of this intervention. Both Briesch and Chafouleas (2009) and Fantuzzo et al. (1987) investigated this topic and found that the total number of student-directed components did not significantly affect outcomes. However, a comparison between intervention protocols has not been conducted to determine the relative effects of student involvement for SM intervention protocols with similar components. Given the variability within the literature on this feature of SM, it is well suited for meta-analytic examination.

Settings. The context of intervention is an important consideration, as variables related to the intervention environment may influence the feasibility of implementation and desired outcomes. SM has been effectively implemented across a variety of environments, including after-school programs (Belfiore, Fritts, & Herman 2008; Cohen, Rubin, & Heinen 1979), residential treatment facilities (Morrow, Burke, & Buell, 1985; Rasing, Coninx, Duker, & van den Hurk 1994), and postsecondary work settings (Lagomarcino & Rusch 1989; Nelson, Lipinski, & Black, 1975; Zegiob, Klukas, & Junginger 1978). However, the school setting is particularly important due to the unique challenges that occur in this context.

To constrain the context of the current investigation, the setting variable for all studies was limited to school settings. Interventions that can be effectively implemented within the school context are of importance, primarily because children ages 5-18 spend a majority of their day in school settings. In addition, classrooms are often home to students with a variety of behavioral needs. In the past, students with behavioral problems were excluded from many educational opportunities. However, given current federal mandates under the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) and the No Child Left Behind Act (NCLB, 2001), educators must consider the least restrictive environment and provide students access to grade-appropriate curricula to the maximum extent possible.

The current study examined the differences between unique school settings with varying levels of instructional support (e.g., self-contained vs. general education).

Student characteristics. In addition to contextual variables, current EBP paradigms call for investigation of SM interventions for specific populations. Student characteristics and outcome variables are important to consider when applying SM interventions (Reid, 1996). Thus, in school settings, it is important to determine if there is a differential effect for SM based on student characteristics such as age, gender, and disability category. While it is important to determine who benefits from SM interventions, it is also important to consider what specific behaviors show reaction to the intervention. Given the widespread application of SM, more information is needed to show what types of students and which behavior problems are most reactive to SM interventions.

A recent meta-analysis of SM interventions for behavior found no differential effects based on gender, age, setting, disability status, or target behavior (Briesch & Chafouleas, 2009). While no significant differences were found on these variables, this meta-analysis evaluated treatment effects through two effect sizes (Cohen's *d* [1988]and PND [percentage of nonoverlapping data]), both of which are questionable methodological choices given the widely noted issues in evaluating treatment effects in single-case research (SCR; Parker et al., 2005). Cohen's *d* is not recommended for application to SCR data due to scale dependency and problematic interpretation (Campbell, 2004). Application of Cohen's *d* to SCR data is also confounded by the fact that SCR data typically do not meet the parametric data assumptions required to apply this statistic (Parker et al., 2005; Parker & Hagan-Burke, 2007).

Research on effect size indices for SCR have called into question the use of PND due to overly conservative measurement (Marquis et al., 2000) and the lack of a known sampling distribution that precludes application of confidence intervals to a PND point estimate (Allison & Gorman, 1993; Parker, Vannest, & Brown, 2009). As a result, this analysis is worth reexamination due to issues noted in previous analyses related to effect size (ES) calculation (Parker, Vannest, & Davis, 2011). Combined with the widespread use of the strategy and the potential importance of this information for future applications of SM, reevaluating the analysis of student characteristic variables to substantiate prior findings would be optimal.

Student outcomes. SM studies have demonstrated efficacy with a variety of behavioral targets. As a behavioral intervention, SM shows promise for implementation with school-age children. In particular, SM has been found to be effective for increasing on-task behavior (DiGangi, Maag, & Rutherford, 1991; Ganz & Sigafoos, 2005; Rooney & Hallahan, 1988) and appropriate conversation (Newman, Buffington, & Hemmes, 1996; Newman, Reinecke, & Meinberg, 2000), and for enhancing social skills (Reinecke, Newman, & Meinberg, 1999; Shearer, Kohler, Buchan, & McCullough 1996; Strain et al., 1994). SM is also effective in decreasing disruptive behavior (Barry & Messer, 2003; Freeman & Dexter-Mazza, 2004; Smith et al., 1988). Despite compelling evidence of positive behavioral outcomes in individual studies, to date, results across SCR studies have only been examined in one previous study (Briesch & Chafouleas, 2009). Given the methodological concerns with this study, a revaluation of this variable is appropriate as well.

In addition to student characteristics and outcomes, several other methodological considerations have been debated in the literature as necessary for SM. The most prominent of these are discussed below.

Cueing. Cueing, in this context, is the act of prompting the student to perform the task of self-assessment and/or self-recording. While some view cueing as a necessary element of the SM protocol (Hallahan & Sapona, 1983; Heins, Lloyd, & Hallahan, 1986), others have questioned the appropriateness of external influences on student's ability to learn self-regulatory behaviors (Snider, 1987). Within the available literature, some studies have employed external cues, including teacher-initiated verbal cues (Agran et al., 2005; Barry & Messer, 2003), audio cues from a tape recorder (Boyle & Hughes, 1994; Brooks, Todd, Tofflemoyer, & Horner, 2003), tactile cues (Amato-Zech et al., 2006; Bowers, Clement, Fantuzzo, & Sorenson, 1985); whereas, others have not cued the student at all (Clees, 1994; Coogan, Kehle, Bray, & Chafouleas, 2007).

Cueing has been examined explicitly by Heins et al. (1986), who found that SM with cueing produced superior intervention effects compared to SM without cueing for task completion behaviors. Despite evidence that cueing procedures have the potential to impact study effects, the specific method of cueing has not been fully examined to determine if differences exist between studies based on the type of cueing methodology used. This variable holds important implications for practice in terms of teacher time and resources.

Reinforcement. Positive reinforcement is a powerful agent in behavior change of school-age children (Cameron & Pierce 1994). Within the SM literature, contingent reinforcement is often used to promote positive behavior change (Cavalier, Feretti, & Hodges, 1997; Crawley, Lynch, & Vannest, 2006; Dalton, Martella, & Marchand-Martella, 1999; Dunlap et al., 1995). However, some researchers have asserted that external reinforcers are not necessary for treatment gains (Snider, 1987). In fact, several studies have effectively implemented SM without any form of reinforcement (Amato-Zech et al., 2006; Guresko-Moore, DuPaul, & White, 2007; Hughes & Hendrickson, 1987; Levendoski & Cartledge, 2000).

Despite conflicting evidence for the use contingent reinforcement with this intervention, the relative effect of reinforcement within a SM intervention package is still not clear. The current analysis will contrast effects between studies based on the use of contingent reinforcement.

Accuracy. The SM literature differs widely in the emphasis placed on student accuracy in recording. Several studies have actively targeted accuracy and offered reinforcement based on student-teacher agreement (Lipinski & Nelson, 1974; Lloyd & Hilliard, 1989; Nelson & Hayes, 1981). Others have indicated positive effects with no attention to the accuracy of student recording (Amato-Zech et al., 2006; Guresko-Moore et al., 2007; Holman & Baer, 1979).

As a result, questions have been raised about the necessity of assessing accuracy in recording. Several studies have shown that high levels of accuracy are not a determining factor for positive treatment gains (Broden et al., 1971; Hughes et al., 1989; Lipinski & Nelson, 1974; Nelson & Hayes, 1981). Given the lack of consensus on the issue of accuracy, an investigation is warranted to compare the relative effects of this variable.

Considering the wide application of SM in behavioral research, additional analysis is needed to determine which components contribute to the success of SM.

Current evaluation paradigms for EBP emphasize the need for clear protocols, giving credence to interventions that have prescriptive steps that guide implementation (Baer et al., 1968; Chambless & Ollendick, 2001; Horner & Kratochwill, 2012).

Research Questions

Based on this survey of available research on factors related to SM, the following research questions were addressed in the current investigation:

1) Does sufficient evidence exists to classify SM interventions as an EBP?
2) Are there differences (in effect) between sets of intervention components?
3) Are there differences within sets of intervention components based on levels of student involvement?
4) Is SM differentially effective based on targeted outcome?
5) Is SM differentially effective based on participant variables (i.e., age, gender, and disability category)?
6) Is SM differentially effective based on student instructional setting?
7) Is SM differentially effective based cueing strategies?
8) Does contingent reinforcement of behavior improve outcomes?
9) Does accuracy of student recording relate to improved behavioral outcomes?

CHAPTER III

METHODS

Research Design

Article selection criteria. A comprehensive literature review was conducted using standard methods identified by Lipsey and Wilson (2001), including (a) reviewing published journal articles on the topic of SM interventions, (b) reviewing references within identified studies (i.e., historical search), and (c) conducting keyword searches from bibliographic databases.

Published studies that conducted SM interventions for behavior with school-aged students were targeted. The following procedures were used to locate articles. First, databases such as the Education Resource Information Center (ERIC), Academic Search Complete (EBSCO), PsychINFO (Proquest), and Cambridge Scientific Abstracts Database were searched for relevant articles. The following keywords were used in the search: *self-monitoring, self-instruction, self-recording, self-evaluation, self-management, self-reinforcement*, *self-observation*, and *self-graphing*. Due to a large overlap of SM studies in disciplines outside of education, search strings were generated by combining keywords, *special education, education, classroom intervention, school*, and *teacher*, with Boolean operators AND, OR added to each of the initial search terms listed above. For example, *self-recording* might be combined with *classroom intervention* for a search term, *self-recording AND classroom intervention.* Second, a hand search was performed by checking the citations from relevant studies to determine if any of the articles cited would qualify for inclusion in the present review.

Following this initial literature search process, a pool of 6,592 possible studies was located. After reading the title and abstract of each of the identified articles, the number of included studies was reduced to 399.

Following the initial search, additional inclusion criteria were applied. Studies were included in the current analysis if the following criteria were met:

1. Utilized SCR methodology with a clearly readable graph of data. Group studies were omitted to allow for continuity in comparison of effect sizes (Lipsey & Wilson, 2001).

2. Clearly identified a behavioral outcome variable. Studies that examined specific academic skill attainment or work completion were omitted. Studies that examined both academic and behavioral dependent variables separately were included; however, only the behavioral outcomes were considered in analysis.

3. Occurred within a school setting. Studies that occurred in residential treatment facilities, hospitals, clinics, homes, private schools, Head Start, or Easter Seals preschool programs were excluded. In studies that examined outcomes across school and other settings, dependent measures from non-school settings were excluded from analysis.

4. Included children or adolescents between the ages of 5 and 21 who were receiving services in school settings.

5. Were peer-reviewed, original reports of experimental research available in English.

6. Showed manipulation of an independent behavioral treatment variable (i.e., self-monitoring and measurement of a behavioral dependent variable).

7. Meet minimum SCR design requirements (Horner et al., 2005; Kratochwill et al., 2010) to demonstrate experimental control on the dependent variable (see discussion under Assessment of Methodological Quality below).

8. Examined SM intervention data in a phase immediately preceded by a baseline or nonexperimental condition phase. Studies that examined multiple intervention protocols (e.g., token economy and SM) were included if the SM intervention was evaluated in a phase adjacent to a baseline phase. The single exception to this criterion applied to studies that collected data in a student training phase between the baseline and intervention phases.

In addition, studies that met the above conditions were excluded in the following cases:

1. Studies of medical outcomes (e.g., diabetes management) or physical performance outcomes (e.g., swimming stroke improvement, golf, or dancing).

2. Data from individuals not targeted by the intervention. Some studies included data for students that were not involved in the intervention. For example, Sainato, Goldstien, and Strain (1992) examined the use of facilitative communication strategies for student peers working with students with an autism spectrum disorder (ASD). The SM intervention was only implemented with the student peers; however, social behaviors were also measured as a secondary outcome with the students with ASD. Since the SM intervention was only directed toward the general education peer working with the

student with ASD, only the data from students directly using the SM intervention were included in the current analysis.

Assessment of methodological quality. In meta-analysis, it is important to use only studies that demonstrate experimental control of the dependent variable (Lipsey & Wilson, 2001). To verify the presence of internal validity for purposes of inclusion, two graduate students with experience and training in research methodology reviewed the methods and data section of each of the included articles. Each of the students coded the results separately, and then compared assessment results. When the student disagreed on a particular study, both would review the article a second time and discuss to consensus.

Within the pool of studies targeted, three designs paradigms were used most often: multiple baseline design (MBD) between subjects or behaviors, single baseline designs (SBD), such as reversal designs and (c) changing criteria, and changing criterion designs. Evaluation procedures for each of these designs are as follows. The "points" of change were evaluated in MBDs as a phase change within a single participant. Therefore, a MBD across three participants with a single phase change (A-B) would be counted as having one point of change for each participant, giving the design a total of three. Within this criterion, the number of participants was an important consideration for determining the level of experimental control. Thus, MBDs with three points of change were included in the analysis because the design was sufficient to demonstrate experimental control according to criteria set by Horner et al. (2005). For studies with a SBD or changing criteria, the number of phase changes was also used to determine the level of experimental control. Therefore, reversal and changing-criteria designs were evaluated to determine if three experimental "points" of control were present. Only studies with three points of control were included in further analysis.

In addition to the assessment of internal validity, the presence of sufficient data and reliability were evaluated. The researcher counted the number of data points in each phase analyzed. Designs that included phases with less than three data points were excluded in the analysis. Reliability was coded, and only studies with acceptable levels of reliability were included in the analysis. Acceptable reliability was set at a minimum of .80 for percent agreement and .60 for Cohen's Kappa (Kratochwill et al., 2010).

A full review of the studies resulted in exclusion of additional articles for the following reasons: 40 studies were not empirical research, 77 did not employ SCR, 43 were eliminated for not meeting minimum design quality standards for SCR, 58 studies examined participants outside of the school setting, 57 studies did not employ an SM intervention, 25 studies did not examine an SM intervention in a phase adjacent to a baseline, and 5 studies included illegible graphs. Application of the additional exclusion criteria resulted in a total of 94 studies considered for further analysis.

Data extraction. Graphic data from published studies was digitized using the GetData digital ruler (GetData, 2012). Digitizing data results in exact reconstruction of the original graphic data to numeric data. For this process, each graph was extracted from PDF versions of the published articles using the "Snipping Tool" from Microsoft Office, Version 2010. This digital snapshot tool allows users to capture just the image of the graphic data to then be uploaded into the GetData program.

Each graph was extracted and labeled separately for each of the included studies. The graphic data were then uploaded into the GetData program where the scale of the x and y axes are set in accordance with information from the graph. Following this procedure, each data point was processed to ensure exact concordance with original study data. Values from the GetData output were rounded to whole numbers whenever necessary to ensure an appropriate match with original study data. Following this digitizing procedure, each data set was entered into an Excel spreadsheet and attached to the variables of interest (moderators, outcomes, etc.) from each study.

Intervention method classification. In the current literature, 11 components may be present under the label SM (Fantuzzo et al., 1987). Each of the intervention components was coded to designate the presence of that component. In addition, student participation within each component was coded to determine the extent to which the student was involved in each component of the intervention (see Table 3).

The researcher coded the presence or absence of each intervention component along with information regarding student involvement or implementation responsibility. If the student was responsible for the component implementation, it was coded with an "S." If a teacher, researcher, or other person was responsible for implementation, it was coded with an "R." This coding strategy was used to assist in determining how important certain components were to overall effects and what impact student involvement had on outcomes for studies that used these components.

Table 3

Component Coding by Presence and Responsibility

Intervention component	Code		
	Teacher/ researcher (R)	Student (S)	Not present (-)
Selection of dependent variable	If a target behavior exists in the study and the student is not explicitly involved in the selection process, then the study is coded “R.”	The study is coded “S” if the article explicitly states that the student is involved in the selection of the target behavior.	Studies without a dependent variable were excluded from analysis.
Definition of target behavior	If a target behavior is defined in the study and the student is not explicitly involved in this process, then this study is coded “R.”	The study is coded “S” if the article explicitly states that the student participates in defining the target behavior.	Studies without a defined dependent variable were excluded from analysis.
Determination of performance goal	The study is coded “R” if the article explicitly states that determination of the performance goal occurs, but the student is not involved in the process.	The study is coded “S” if the article explicitly states that the student is involved in determining a defined performance criterion for the target behavior.	If no performance goal is explicitly defined, then the study is coded “-” for this variable.
Evaluation to determine whether performance goal was met	The study is coded “R” if the article explicitly states that evaluation of the performance goal occurs, but the student is not involved in the process.	The study is coded “S” if the article explicitly states that the student is directly involved in making the comparison to determine if the goal was met.	If a performance goal is not evaluated, then the study is coded “-” for this variable.
Instructional prompts for target behavior	The study is coded “R” if the article explicitly states that someone other than the student delivers an explicit instructional prompt during the intervention.	The study is coded “S” if the article explicitly states that the student delivers explicit instructional prompts during the intervention.	If instructional prompts are not present, then the study is coded “-” for this variable.

Table 3 (Continued)

Component Coding by Presence and Responsibility

Intervention component	Code		
	Teacher/ researcher (R)	Student (S)	Not present (-)
Observation of target behavior	The study is coded “R” if the article explicitly states that someone other than the student judges whether or not the target behavior occurs.	The study is coded “S” if the article explicitly states that the student judges whether or not the target behavior occurs.	Studies where no judgment occurs regarding the presence of the target behavior were excluded from analysis.
Recording	The study is coded “R” if the article explicitly states that someone other than the student records whether or not the target behavior occurs.	The study is coded “S” if the article explicitly states that the student records whether or not the target behavior occurs.	Studies where no record of the occurrence of the behavior occurs were excluded from analysis.
Graphing or charting behavior	The study is coded “R” if the article explicitly states that someone other than the student summarizes the recording of behavior as part of the intervention.	The study is coded “S” if the article explicitly states that the student summarizes the recording of behavior either graphically or with written notes.	Studies that do not summarize behavioral data within the intervention are coded “-” for this variable.
Selection of primary reinforcer	The study is coded “R” if external reinforcement is present, but the student is not involved in the selection.	The study is coded “S” if the article explicitly states that the student is involved in the selection of the primary external reinforcer.	Studies with no external reinforcement are coded “-” for this variable.
Administration of secondary reinforce	The study is coded “R” if secondary reinforcement is present, but the student is not involved in the administration.	The study is coded “S” if the article explicitly states that the student is involved in the administration of secondary reinforcers.	Studies with no secondary reinforcement are coded “-” for this variable.

Table 3 (Continued)

Component Coding by Presence and Responsibility

Intervention component	Code		
	Teacher/ researcher (R)	Student (S)	Not present (-)
Administration of primary reinforce	The study is coded "R" if primary external reinforcement is present, but the student is not involved in the decision to administer.	The study is coded "S" if the article explicitly states that the student is involved in the decision to administer.	Studies with no primary external reinforcement are coded "-" for this variable.

Note. Intervention component definitions and criteria adapted from Fantuzzo et al. (1978) and Briesch and Chafouleas (2009).

Evaluation of Interventions as Evidence Based Practice

Competing paradigms exist across fields for classifying an intervention as an EBP (Gillam & Gillam, 2006; Tate et al., 2008). One set of evaluation procedures has emerged as the most prominent for evaluating SCR in educational research (Horner & Kratochwill, 2012). In addition to minimum design requirements, several other standards must be met within this paradigm. Specifically, an intervention protocol must be supported by a minimum of five studies that document adequate experimental control. Studies must be conducted by at least three research groups in three separate geographic locations. Finally, studies must include effects for 20 separate participants.

Given the emphasis on clear intervention protocols, the above criteria were applied to groups of studies with like components. Following the coding of each study with the Fantuzzo et al. (1987) framework for identifying intervention components, studies with identical component packages were grouped and evaluated. For example, all of the studies with Recording, Self-instruction, and Self assessment were grouped and evaluated with the above criteria.

Given the volume of combinations of components anticipated, each unique combination of intervention components will be coded with a unique Alpha code for the evaluation of intervention component presence. To identify intervention packages based on differences in student responsibility, each unique package of intervention components will be identified with an Alpha plus a number. The total number will be determined by the number of unique combinations of components within each major category.

Coding of geographic regions. To determine that an intervention meets EBP standards, research must be conducted in a minimum of three geographic regions (Horner & Kratochwill, 2012). Regional designations were made using the United States Census Bureau (USCB) regional designation system for the university where the first author published the study (United States Census Bureau, 2012). This regional coding system was selected due to the fact that this method of designating regions has remained unchanged in the United States since the 1970s (United States Census Bureau, 2012). The USCB regional designation system divides the United States into nine separate regions based on population density (see Table 4). For studies authored in countries outside of the United States, each country was counted as a separate region.

Table 4

Geographic Regions

Division	States
1	Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut
2	New York, Pennsylvania, New Jersey
3	Wisconsin, Michigan, Illinois, Indiana, Ohio
4	Missouri, North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa
5	Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida
6	Kentucky, Tennessee, Mississippi, Alabama
7	Oklahoma, Texas, Arkansas, Louisiana
8	Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, New Mexico
9	Alaska, Washington, Oregon, California, Hawaii

Note. Information contained in this table reprinted from the U.S. Census Bureau. http://www.census.gov/geo/www/us_regdiv.pdf

Additional moderating variables. A moderator is an independent variable that affects the primary relationship between the intervention and behavior change outcome (Baron & Kenny, 1986). Each of the studies was coded on moderating variables identified in the literature as critical to the implementation of SM interventions. These variables were as follows: Student Characteristic Variables, Setting, Outcome variables, Cueing, Reinforcement, and Accuracy. The operational definitions and coding procedures used for each of these potential moderator variables are discussed below.

Student characteristics. The Student Characteristic Variables analyzed consisted of student age, gender, and disability category. The age variable had two levels: students in primary settings vs. secondary settings. The information provided for this variable was not consistent among the studies. Some studies reported only grade, while other reported ages. Therefore, students in the primary category were defined as prekindergarten to 6th grade or 3-12 years old. The secondary category was defined as students in 7th-12th grade or 13-21 years old. Gender was defined as male or female.

The Disability category was defined according to federal classification standards for education. Disability categories from the Individuals with Disabilities Education Act (IDEA, 1997) or the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) were coded according to special education eligibility categories provided by each study's authors. Students with clinical diagnoses were grouped into an IDEIA category when sufficient details on the diagnosis were available in the study text. For example, a subject diagnosed by a physician with Major Depressive Disorder would be coded as Emotionally Disturbed or students listed as having a Developmental Disorder, Pervasive Developmental Disorder, or Asperger Syndrome were coded under the umbrella term Autism. Participants with multiple diagnoses were only coded based on their primary disability category.

Participants who were not identified as receiving special education services and demonstrated significant behavior problems were identified as At Risk. Participants who were not identified as receiving special education services and demonstrated no significant behavior problems were classified as Not Disabled.

This coding scheme in not ideal; due to the potential differential impact of comorbid conditions, all possible combinations of disabilities were not evaluated. While the number of participants in the current study is large, the number of participants needed to evaluate all possible disabling combinations is beyond the limits of the current study or the existing literature base for this intervention.

Setting. The Instructional Setting category was coded based on the student's educational placement and educational support within the classroom. The instructional setting variable was only considered for the setting in which the intervention was applied. For example, a secondary student may have access to a general education physical education class with no special education teacher support; however, the intervention was applied within a self-contained reading classroom with multiple teachers and instructional aides. For this condition, the study would be coded as Self-Contained. This setting variable in the current study had three levels, Mainstream (fully educated in the general education classroom with no instructional or behavior support beyond the single teacher), Inclusion (educated in the mainstream classroom with additional instructional or behavior supports), and Self-Contained (educated outside the mainstream classroom with instructional or behavior support). Instructional or behavioral support in the current study was defined as a teacher, researcher, or an instructional aide placed within the classroom to support student curricular or behavioral needs.

Outcomes. In the investigation of the effectiveness of SM for behavior, several dependent variables were identified in published articles. All were collapsed into five categories that captured similar behaviors under a common label: Disruptive, Functional Communication, Social Behavior, On Task, and Following Rules. The Disruptive category included behaviors that are distracting to others in the classroom – both verbal and nonverbal. The Disruptive category included behaviors such as, talking out, yelling, screaming, out of seat, and aggression. The Social Behavior category included social interaction outcomes that are not strictly communicative in nature, such as sharing, positive interactions, social skill improvement, and use of social facilitation strategies. The Functional Communication category examined both verbal and nonverbal communication outcomes. This included verbal behavior outcomes such as requesting, initiating verbalizations, and appropriate commenting. This category also included nonverbal behaviors such as appropriate eye contact and raising head to the appropriate position to communicate. The On Task category examined student attention to presented tasks and student engagement. The On Task variable did not include task completion outcomes. Data from task completion outcomes were excluded from the current study. Of note, five studies aggregated On Task and Disruptive behaviors as a single outcome variable. Results from these studies are aggregated separately from both the studies in the On Task and Disruptive categories. Finally, the Following Rules category included desirable classroom behaviors that were not strictly disruptive, social, or communicative in nature. This included several discreet outcomes such as classroom work preparation, following teacher directions, and transitioning appropriately. This category also included student outcomes in studies that aggregated classroom rule sets. The Following Rules category was not fully independent from the other categories, as many of the rule sets included on task, nondisruptive, and social behaviors among other targeted outcomes. Due to the aggregation of varying outcomes in some studies, this category functioned more as a general measure of SM effectiveness in classrooms rather than an indicator of specific behavioral outcomes. Despite the lack of specificity in the Following Rules category, the creation of this category allowed for the expression of all study outcomes and preserved the integrity of each of the four categories mentioned above.

Separate from student characteristics and dependent variables, several intervention procedures or independent variables were examined to determine differences in study effects.

Cueing. This variable was based on the presence and form of external prompting for the student to engage in the SM intervention. There were seven levels of this moderator: (a) teacher-initiated cue, (b) auditory cue, (c) tactile cue, (d) visual cue, (e) peer cueing, (f) time cue, and (g) no cueing. Studies were coded as Teacher-Initiated Cue if the teacher or another adult was responsible for prompting the student to engage in the intervention. Studies were coded as Auditory cue if students were prompted to engage in the intervention by an auditory cue such as a bell, alarm, or tones played by an audio device. This variable was separate from the Teacher Initiated Cue in that prompting did not involve the teacher or another adult speaking. Studies were coded as Tactile cue if a vibro-tactile device (e.g., the MotivAider®) was used without additional input from the teacher or other adults. Studies were coded as a Visual cue, if the method of cueing was exclusively visual in nature. Examples of this type of cueing include prompting cards with words and/or pictures that the student manages. Studies were coded as Peer cue if a student peer was responsible for prompting the target student to engage in the intervention. Studies were coded as Time cue if the cue to engage in SM was at the completion of the activity or class period. Finally, studies were coded as No Cue if there was no external prompt for the student to engage in the intervention.

Reinforcement. The Reinforcement variable was based on the presence and form of reinforcing the student. There are two occasions where reinforcement is commonly used in SM, reinforcement of improvement on the outcome variable and reinforcement of accuracy in recording. This variable examined the presence or absence of a reinforcer or reinforcement condition only in relation to the improvement of behavior. Reinforcement for accuracy in recording was included in the Accuracy variable. The Reinforcement moderator was coded at two levels: No Reinforcement and Contingent Reinforcement for behavior improvement. The Contingent Reinforcement variable was operationalized as any study that implemented a contingent reinforcement for improvement on the outcome variable or adherence to intervention implementation. The absence of contingent reinforcement was coded as No Reinforcement.

Accuracy. This variable was based on the presence or absence of external checks of student accuracy in recording. Studies were coded as employing an accuracy check if ongoing accuracy of student recording was directly assessed within the study. This moderator was coded at three levels: No Accuracy Check, Accuracy Checked (no contingent reinforcement; many studies trained students to become accurate in self-recording as part of the initial intervention training; however, only studies that utilized rating accuracy checks as a part of the intervention –that is, post training – were coded with this variable), and Accuracy Checked (contingent reinforcement – if student accuracy in recording was tied to contingent reinforcement).

Reliability of coding. To assess the reliability of data coding, a doctoral student in special education who had not participated in the original coding and was blind to previous coding results recoded each variable for 20% of the studies analyzed. These results were compared to the original data coding. Reliability was calculated using a simple percent agreement or (total agreement/agreement + disagreement). Initial agreement was 87%.

Cohen's Kappa (Kappa) was also calculated. Kappa is a more conservative measure of reliability that adjusts for expected chance agreement (Ary & Suen, 1989). Initial Kappa was an acceptable 74%; Kappa values above 60% are considered good agreement (Altman, 1991).

Following this initial assessment of reliability, the graduate students discussed each of the variables in an attempt to come to consensus. Percent agreement was 100% and Kappa was 100%. Reliability under all four of these conditions was above acceptable limits.

Data Analyses

Phase contrast selection. Selecting which phase contrasts to evaluate is an important consideration to protecting the integrity of results. Only phase contrasts that represented independent manipulation of the independent variable were evaluated with an ES. This resulted in the forward evaluation of any adjacent baseline-to-intervention phases. Data from subsequent intervention phases were not aggregated with a prior intervention phase. Each phase and phase combination in the design was only evaluated once to preserve the independence of all contrasts. For example, for designs that employed reversal logic, separate effect sizes were calculated for each baseline/intervention combination. Each of the separate effect sizes in this case was aggregated to reflect the overall outcome on the dependent variable. Therefore, an ABABAB design produced three separate effect sizes, which were then aggregated into one omnibus effect size for the design.

MBDs were treated with similar logic. In the current application, given an appropriate baseline and intervention phase, each tier of the MBD was evaluated separately for effect, and then these ESs were aggregated using the methods described below.

Effect size. For the current study, the Tau-U ES was used to determine intervention effects. Tau-U is a method for measuring data nonoverlap between two phases (A and B). It is a “distribution-free” nonparametric technique, with statistical power of 91% to 95% of (ordinary least squares, OLS) linear regression when data conform to parametric assumptions. When data do not conform to parametric data assumptions, which is common in SCR, the power of a nonparametric statistic can exceed the parametric statistical analogue (Cliff, 1993; Delaney & Vargha, 2002; Wilcox, 2010). As a result, this is an index that is well suited for small datasets.

Tau-U follows the “S” sampling distribution (Parker, Vannest, Davis, & Sauber, 2011), making it possible to calculate exact *p*-values and confidence intervals. Tau-U may be interpreted as the “percent of data that improve over time.” Tau-U analysis yields scores between -1.0 and 1.0, with a score of 0 indicating no difference between phases. Scores above 0 indicate improved performance across phases. Conversely, scores below 0 indicate deterioration in performance (Parker et al., 2011). Tau-U scores from individual phase contrasts can be aggregated to provide a single omnibus ES for a variety of SCR designs, individual phase contrasts. Tau-U is also useful for a range of simple to complex designs.

Effect size aggregation. ESs from available studies were combined to determine omnibus effects, in addition to differences between intervention component sets and moderators. Tau-U was aggregated using similar methods and presented separately.

The Tau-U effect size is particularly innovative because multiple phase contrasts can be easily aggregated. Tau-U uses the S distribution to determine the variance score (Var_s). Tau-U effects were averaged after weighting each ES by the inverse of the variance score (Var_s). Furthermore, Tau-U standard errors were aggregated by the following methods: A varience is associated with each Tau-U. The variance of the aggregated Tau-U is the inverse of the inverse sum of the individual variances. The standard error of the aggregated Tau-U is then equivalent to the square root of the aggregated variance. For example, if two studies are to be aggregated with Tau U's τ_1 and τ_2 and associated variances v_1 and v_2, then the aggregated Tau-U is calculated with the following formulas (see Figure 1 & 2)

$$\tau_a = \frac{\frac{1}{v_1}\tau_1 + \frac{1}{v_2}\tau_2}{\frac{1}{v_1} + \frac{1}{v_2}}$$

Figure 1. Formula for aggregating Tau-U values

The variance of the aggregated τ_a is then calculated with the following formula:

$$v_a = \frac{1}{\frac{1}{v_1} + \frac{1}{v_2}}$$

Figure 2. Formula for the aggregated Tau-U variance

Comparing effects. Analysis of intervention components and moderators followed standard practice for analyzing categorical variables (Agresti, 2010; Siegel & Castellan, 1988). Statistical significance for moderator variables with two groups was evaluated using the Wilcoxon signed-ranks statistic (Wilcoxon, 1945). Moderator variables with three or more groups were evaluated with the Kruskal-Wallis one-way analysis of variance (Kruskal & Wallis, 1952). In cases where the Kruskal-Wallis showed significant differences within groups of variables, the Dunn post-hoc test (Dunn, 1964) was used to determine significance between each pairwise combination of groups. The Dunn post-hoc test is a nonparametric method for comparing pairwise differences between groups. As such, it is the recommended method for evaluating data that (a) do not meet the normal distribution assumption and (b) have unequal samples sizes (Hollander & Wolfe, 1999).

Effect size calculation. Effect size calculation and aggregation were analyzed using original software developed by the author using the Maple platform (Maplesoft, 2012). The Kruskal-Wallis and Dunn post-hoc test were analyzed with SAS (Version 9.3) statistical software.

Levels of Component Analysis

Two levels of analysis were necessary to answer research questions 1-3; specifically, (a) to determine if continuity exists between intervention components for purposes of determining which SM interventions meet EBP guidelines, (b) to determine if effects differ based on intervention components, and (c) to determine if there are differences among sets of intervention components based on levels of student involvement.

First level. Effects for SM were calculated based on the presence of only the specific intervention components assigned to each of the intervention methods. This analysis was used to determine how many studies with similar methods aligned. The presence of the component was the only factor used to include studies in each of the intervention analysis. Studies were aggregated based on the presence of like components. Effect sizes were calculated for each of these groups based on the presence of only the specific intervention components. This analysis was used to determine how many studies with similar methods aligned within the broader SM construct. The presence and absence of individual components were the only factors used to include studies in each analysis. This analysis step allowed for the examination of intervention components separate from overarching intervention category.

Second level. At the second level of analysis, implementation responsibility was considered. Within each of the intervention categories determined in the first level of analysis, studies were aggregated based on the use of similar methods for student implementation. Clustering studies in this manner allowed for partitioning effects based on the degree of student responsibility for SM implementation.

Combinations of student-implemented components within each intervention category were analyzed to determine how student involvement affected the outcome for each main intervention category. For example, a SM intervention with three components (e.g. observation of behavior, recording, and graphing), has six possible combinations of each of these components. Each possible combination of student-implemented components was analyzed separately to determine its relative effect for student involvement. Since study methods could not overlap for any data array, each of the effect sizes calculated showed an independent estimate of effect based on common study methods used.

The aggregate effect for studies with these categories allowed comparison of the relative contribution of each discrete component. This method of coding differed from past analysis, in that each component remained discrete in the analysis. Past analysis of this variable (Briesch & Chafouleas, 2009; Fantuzzo & Polite, 1990) has examined the total number of intervention components. This aggregate method likely masked the relative effects of certain intervention components. Therefore, a more careful treatment of this variable was considered necessary.

CHAPTER IV

RESULTS

Data from this study yielded 655 separate effect sizes from 94 unique studies with 288 participants. The omnibus Tau-U across all SM studies was .77 CI_{95} [.76, .79]. Within these studies, a broad range of Tau-U values were identified (from -0.20 to 1.00). Given the broad range of ES across studies, additional analyses were conducted to answer questions that are critical to the implementation of SM.

Research Question 1. Does Sufficient Evidence Exist to Classify SM Interventions as an EBP?

Within the SM literature, several groups of interventions emerged that meet EBP guidelines. After applying the Fantuzzo et al. (1990) classification rubric to all studies included in the analysis, four unique SM intervention protocols emerged as meeting current EBP guidelines in education (see Table 5). These four intervention protocols ranged in number of total intervention procedures from four to nine. Two additional intervention protocols approached criteria for EBP, as intervention packages M and P met criteria in every area except for the minimum number of subjects (e.g. 20).

Table 5

Outcomes by Component Presence

Int. package marker	Components											Tau-U	95% CI	# of studies	# of subjects	# of ESs	# of unique geo regions
	Select DV	Def DV	Obs.	Rec.	Inst. pro.	Det. goal	Eval. goal	Sel. rein	Admin sec. Rein	Admin pri. Rein	Graph/ chart						
A**	X	X	X	X	-	-	-	-	-	-	-	.82	.77 - .86	20	50	99	9
B	X	X	X	X	-	-	-	-	-	-	X	.38	.30 - .46	1	7	28	1
C	X	X	X	X	-	-	-	X	X	X	-	.72	.61 - .83	3	13	19	2
D**	X	X	X	X	X	-	-	-	-	-	-	.81	.77 - .86	14	53	75	6
E	X	X	X	X	X	-	-	-	-	-	X	.82	.67 - .96	4	11	11	3
F**	X	X	X	X	X	X	X	-	-	-	-	.89	.84 - .95	6	24	65	4
G	X	X	X	X	-	X	X	-	-	-	-	.93	.80 - 1.00	3	13	15	3
H	X	X	X	X	-	-	-	X	-	X	-	.65	.54 - .76	4	10	15	4
I	X	X	X	X	X	-	-	X	X	X	-	.39	.33 - .46	2	8	14	2
J	X	X	X	X	X	-	-	X	X	X	X	.82	.64 - .99	1	4	10	1
K	X	X	X	X	X	X	X	X	-	X	-	.83	.75 - .90	2	6	58	2
L**	X	X	X	X	-	X	X	X	X	X	-	.76	.73 - .79	16	45	127	9
M†	X	X	X	X	-	X	X	X	-	X	-	.96	.90 - 1.00	6	14	43	6
N	X	X	X	X	-	X	X	X	-	X	X	.74	.57 - .91	2	4	5	2
O	X	X	X	X	-	X	X	X	X	X	X	.79	.70 - .89	4	9	20	3
P	X	X	X	X	X	X	X	X	-	X	X	.94	.71 - 1.00	1	1	2	1
Q†	X	X	X	X	X	X	X	X	X	X	-	.88	.82 - .95	9	18	47	5
R	X	X	X	X	X	X	X	X	X	X	X	.72	.46 - .99	1	1	2	1

Note. Select DV = Selection of dependent variable; Def DV = Definition of target beh.; Obs = Observation of target beh.; Rec. = Recording of behavior; Inst. pro. = Instructional prompts; Det. goal = Determination of performance goal; Eval goal = Evaluation of performance goal; Admin sec. rein = Administration of secondary reinforcer; Admin pri rein = Administration of primary reinforcer; CI = Confidence interval. ** Meet EBP requirements. † Approach EBP requirements.

Analysis of the four intervention protocols that meet EBP requirements indicated a statistically significant difference among these intervention sets (Kruskal-Wallis $p >$ .000). While SM packages A (Selection of the DV, Defining the DV, Observation and Recording) and D (Selection of the DV, Defining the DV, Observation, Recording, and Instructional Prompts) showed identical Tau-U values (ES = .81). These intervention packages only differed by the addition of instructional prompting in Package D. The Dunn post-hoc procedure indicated a statistically significant difference between SM package F (Selection of the DV, Defining the DV, Observation, Recording, Instructional Prompts, Determining Goals, and Evaluating Goals) and SM package L (Selection of the DV, Defining the DV, Observation, Recording, Determining Goals, Evaluating Goals, Selecting Reinforcers, Administering Primary and Secondary Reinforcers). This difference indicates a potential negative effect for SM interventions that include administration of primary and secondary reinforcers compared to an intervention protocol that includes instructional prompting without administration of reinforcers (see Table 6).

Table 6

Statistical Significance between Intervention Packages That Meet EBP Criteria

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	A-D	9.3413	35.2389	
2	A-F	23.7404	40.4364	
3	A-L	30.3526	30.685	
4	D-F	33.0817	41.6008	
5	D-L	21.0113	32.2039	
6	F-L	54.093	37.8208	**

Research Question 2. Are There Differences (in effect) Between Sets of Intervention Components?

The current analysis found 18 unique intervention packages among the combinations of 11 potential components (see Table 7). Tau-U ESs ranged from .38 CI_{95} [.30, .46] to .96 CI_{95} [.90, 1.00]. A Kruskal-Wallis analysis showed significant differences among treatment packages within these groups (p = <.0001). Examination of statistical significance following the Dunn post-hoc procedure indicated differences between SM package B (Selection of the DV, Defining the DV, Observation, Recording, and Graphing and Charting) and seven other intervention packages: SM package A, D, F, G, K, M, and Q. Similarly, statistically significant differences were found between SM package I (Selection of the DV, Defining the DV, Observation, Recording, Instructional Prompting, Selecting Reinforcers, Administering Primary and Secondary Reinforcers) and four other SM intervention packages: SM package A, F, G, and M. SM packages B and I both showed relatively low aggregate ESs. The ES for these two packages were

Tau-U = .38 CI_{95} [.30, .46] and .39 CI_{95} [.33, .46], respectively. In addition, results from the previous analyses were substantiated, and a statistically significant difference was detected between SM package F and SM package L.

Table 7

Statistical Significance between All Intervention Packages

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	A-B	160	128.257	**
2	A-C	32.595	124.71	
3	A-D	15.336	80.5	
4	A-E	8.201	141.847	
5	A-F	38.763	92.373	
6	A-G	57.638	124.71	
7	A-H	0.708	141.847	
8	A-I	131.786	128.257	**
9	A-J	58.679	223.825	
10	A-K	9.905	109.378	
11	A-L	49.102	70.097	
12	A-M	70.071	104.125	
13	A-N	80.179	223.825	
14	A-O	7.629	147.762	
15	A-P	7.071	437.513	
16	A-Q	3.28	104.125	
17	A-R	96.929	437.513	
18	B-C	127.405	161.31	
19	B-D	144.664	130.186	**
20	B-E	151.799	174.896	
21	B-F	198.763	137.844	**
22	B-G	217.638	161.31	**
23	B-H	160.708	174.896	
24	B-I	28.214	164.067	
25	B-J	101.321	246.101	
26	B-K	169.905	149.772	**
27	B-L	110.898	124.023	

Table 7 (Continued)

Statistical Significance between All Intervention Packages

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
28	B-M	230.071	145.98	**
29	B-N	79.821	246.101	
30	B-O	152.371	179.727	
31	B-P	167.071	449.317	
32	B-Q	163.28	145.98	**
33	B-R	63.071	449.317	
34	C-D	17.259	126.693	
35	C-E	24.394	172.312	
36	C-F	71.358	134.55	
37	C-G	90.233	158.504	
38	C-H	33.303	172.312	
39	C-I	99.19	161.31	
40	C-J	26.083	244.271	
41	C-K	42.5	146.746	
42	C-L	16.507	120.351	
43	C-M	102.667	142.874	
44	C-N	47.583	244.271	
45	C-O	24.967	177.213	
46	C-P	39.667	448.317	
47	C-Q	35.875	142.874	
48	C-R	64.333	448.317	
49	D-E	7.135	143.593	
50	D-F	54.099	95.033	
51	D-G	72.974	126.693	
52	D-H	16.044	143.593	
53	D-I	116.45	130.186	
54	D-J	43.343	224.936	
55	D-K	25.241	111.634	
56	D-L	33.766	73.567	
57	D-M	85.407	106.492	
58	D-N	64.843	224.936	
59	D-O	7.707	149.439	
60	D-P	22.407	438.082	
61	D-Q	18.616	106.492	
62	D-R	81.593	438.082	
63	E-F	46.964	150.571	

Table 7 (Continued)

Statistical Significance between All Intervention Packages

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
64	E-G	65.839	172.312	
65	E-H	8.909	185.093	
66	E-I	123.584	174.896	
67	E-J	50.477	253.449	
68	E-K	18.106	161.562	
69	E-L	40.901	138.03	
70	E-M	78.273	158.053	
71	E-N	71.977	253.449	
72	E-O	0.573	189.664	
73	E-P	15.273	453.383	
74	E-Q	11.481	158.053	
75	E-R	88.727	453.383	
76	F-G	18.875	134.55	
77	F-H	38.055	150.571	
78	F-I	170.548	137.844	**
79	F-J	97.441	229.453	
80	F-K	28.858	120.477	
81	F-L	87.865	86.398	**
82	F-M	31.309	115.728	
83	F-N	118.941	229.453	
84	F-O	46.391	156.156	
85	F-P	31.691	440.418	
86	F-Q	35.483	115.728	
87	F-R	135.691	440.418	
88	G-H	56.93	172.312	
89	G-I	189.424	161.31	**
90	G-J	116.317	244.271	
91	G-K	47.733	146.746	
92	G-L	106.74	120.351	
93	G-M	12.433	142.874	
94	G-N	137.817	244.271	
95	G-O	65.267	177.213	
96	G-P	50.567	448.317	
97	G-Q	54.358	142.874	
98	G-R	154.567	448.317	
99	H-I	132.494	174.896	
100	H-J	59.386	253.449	

Table 7 (Continued)

Statistical Significance between All Intervention Packages

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
101	H-K	9.197	161.562	
102	H-L	49.81	138.03	
103	H-M	69.364	158.053	
104	H-N	80.886	253.449	
105	H-O	8.336	189.664	
106	H-P	6.364	453.383	
107	H-Q	2.572	158.053	
108	H-R	97.636	453.383	
109	I-J	73.107	246.101	
110	I-K	141.69	149.772	
111	I-L	82.684	124.023	
112	I-M	201.857	145.98	**
113	I-N	51.607	246.101	
114	I-O	124.157	179.727	
115	I-P	138.857	449.317	
116	I-Q	135.065	145.98	
117	I-R	34.857	449.317	
118	J-K	68.583	236.811	
119	J-L	9.577	221.426	
120	J-M	128.75	234.431	
121	J-N	21.5	306.942	
122	J-O	51.05	256.806	
123	J-P	65.75	485.317	
124	J-Q	61.958	234.431	
125	J-R	38.25	485.317	
126	K-L	59.007	104.381	
127	K-M	60.167	129.707	
128	K-N	90.083	236.811	
129	K-O	17.533	166.779	
130	K-P	2.833	444.296	
131	K-Q	6.625	129.707	
132	K-R	106.833	444.296	
133	L-M	119.173	98.863	**
134	L-N	31.077	221.426	
135	L-O	41.473	144.102	
136	L-P	56.173	436.29	

Table 7 (Continued)

Statistical Significance between All Intervention Packages

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
137	L-Q	52.382	98.863	
138	L-R	47.827	436.29	
139	M-N	150.25	234.431	
140	M-O	77.7	163.382	
141	M-P	63	443.032	
142	M-Q	66.792	125.308	
143	M-R	167	443.032	
144	N-O	72.55	256.806	
145	N-P	87.25	485.317	
146	N-Q	83.458	234.431	
147	N-R	16.75	485.317	
148	O-P	14.7	455.268	
149	O-Q	10.908	163.382	
150	O-R	89.3	455.268	
151	P-Q	3.792	443.032	
152	P-R	104	613.883	
153	Q-R	100.208	443.032	

Research Question 3. Are There Differences Within Sets of Intervention Components Based on Levels of Student Involvement?

Each of the 18 unique intervention packages identified in the above analysis was analyzed separately to determine if differences occurred based on student implementation responsibility. A total of 44 separate interventions were identified when SM packages were disaggregated based on student involvement in each component (see Table 8). Six of the SM packages showed no variation in implementation responsibility. Therefore, SM packages B, E, I, J, P, and R were not examined.

Within SM package A (Selection of the DV, Defining the DV, Observation and Recording), two variations of the intervention occurred within the literature. There was no significant difference ($p = .41$) between these interventions given higher student involvement in SM package A2 vs. A1. Within SM package C (Selection of the DV, Defining the DV, Observation, Recording, Selecting the Reinforcer, and Administering Primary and Secondary Reinforcers), two variations of the intervention occurred within the literature. There was no significant difference ($p = .20$) between these interventions based given higher student involvement in SM package C2 vs. C1. Within SM package D (Selection of the DV, Defining the DV, Observation, Recording, and Instructional Prompting), two variations of the intervention occurred within the literature. A significant difference was detected between intervention packages (Wilcoxon $p = .001$). Within this analysis, SM packages that allowed the student to record behavioral data rather than the researcher/teacher had improved effects. This finding should be interpreted with caution, however, given the relatively low number of ESs available for SM package D2. Within SM package F (Selection of the DV, Defining the DV, Observation, Recording, Instructional prompting, Determining the Goal and Evaluating Goal Attainment), three variations of the intervention occurred within the literature. SM package F1 was not included in the current analysis because it had fewer than five ESs. Therefore, analysis was conducted on SM package F2 and SM package F3. There was no significant difference (Wilcoxon $p = .83$) between these interventions given higher student involvement in SM package F3 vs. F2. Within SM package G (Selection of the DV, Defining the DV, Observation, Recording, Determining the Goal and Evaluating

Goal Attainment), two variations of the intervention occurred within the literature. However, there was no significant difference (Wilcoxon p = .38) between these interventions given higher student involvement in SM package G1 vs. G2. Within SM package H (Selection of the DV, Defining the DV, Observation, Recording, Selection of Reinforcers, and Administration of Primary Reinforcers), three variations of the intervention occurred within the literature. SM package H3 was not included in the current analysis because it had fewer than five ESs. Therefore, analysis was conducted on SM package H1 and SM package H2. There was no significant difference (Wilcoxon p = .051) between these interventions given higher student involvement in SM package H2 vs. H1. Within SM package K (Selection of the DV, Defining the DV, Observation, Recording, Instructional Prompting, Selection of Reinforcers, Administration of Primary Reinforcers, Determining the Goal and Evaluating Goal Attainment), two variations of the intervention occurred within the literature. However, SM package K2 could not be included in the current analysis because it had fewer than five ESs. Given that there were only two variations of this intervention, no statistical significance testing was conducted within SM package K.

Table 8

Aggregated Results by Student Participation

Int. package marker	Components											Tau-U	95% CI	# of studies	# of subjects	# of ESs	# of unique geo-graphic regions
	Select DV	Def DV	Obs.	Rec.	Inst. pro.	Det. goal	Eval. goal	Sel. Rein	Admin sec. rein	Admin pri. rein	Graphing/ charting						
A1	R	R	S	R	-	-	-	-	-	-	-	.55	.37 - .73	2	5	5	2
A2**	R	R	S	S	-	-	-	-	-	-	-	.83	.78 - .88	18	45	94	8
B	R	R	S	S	-	-	-	-	-	-	S	.38	.30 - .46	1	7	28	1
C1	R	R	S	S	-	-	-	R	S	R	-	.73	.62 - .84	2	12	17	2
C2	R	R	S	S	-	-	-	S	S	R	-	.59	.16 - 1.00	1	1	2	1
D1**	R	R	S	S	S	-	-	-	-	-	-	.84	.79 - .89	13	47	69	6
D2	R	R	S	R	S	-	-	-	-	-	-	.18	-.05 - .41	1	6	6	1
E	R	R	S	S	S	-	-	-	-	-	S	.82	.67 - .96	4	11	11	3
F1	R	R	S	R	S	R	R	-	-	-	-	.94	.55 - 1.00	1	1	2	1
F2	R	R	S	S	S	R	R	-	-	-	-	.96	.87 - 1.00	4	13	27	4
F3	R	R	S	S	S	S	S	-	-	-	-	.84	.76 - .91	2	10	36	1
G1	R	R	S	S	-	S	S	-	-	-	-	.98	.79 - 1.00	1	6	6	1
G2	R	R	S	S	-	R	R	-	-	-	-	.88	.70 - 1.00	2	7	9	2
H1	R	R	S	S	-	-	-	R	-	R	-	.87	.71 - 1.00	2	6	7	2
H2	R	R	S	S	-	-	-	S	-	R	-	.35	.19 - .52	1	3	6	1
H3	R	S	S	S	-	-	-	R	-	R	-	1.00	.61 - 1.00	1	1	2	1
I	R	R	S	S	S	-	-	S	S	R	-	.39	.33 - .46	2	8	14	2
J	R	R	S	S	S	-	-	R	S	R	S	.82	.64 - .99	1	4	10	1
K1	S	S	S	S	S	R	R	S	-	R	-	.84	.76 - .92	1	5	56	1
K2	R	R	S	R	S	R	S	R	-	S	-	.72	.43 - 1.00	1	1	2	1
L1	R	R	S	S	-	R	R	R	S	R	-	.77	.70 - .85	4	16	34	4

Table 8 (Continued)

Aggregated Results by Student Participation

Int. package marker	Components: Select DV	Def DV	Obs.	Rec.	Inst. pro.	Det. goal	Eval. goal	Sel. Rein	Admin sec. rein	Admin pri. rein	Graphing/ charting	Tau-U	95% CI	# of studies	# of subjects	# of ESs	# of unique geo-graphic regions
L2	R	R	S	S	-	S	S	S	S	S	-	.81	.45 - 1.00	1	2	2	1
L3	R	R	S	S	-	R	S	R	S	S	-	.95	.84 - 1.00	1	2	3	1
L4	R	R	S	S	-	R	R	S	S	R	-	.75	.70 - .79	3	10	57	2
L5	R	R	S	S	-	R	S	R	S	R	-	.77	.62 - .92	2	4	9	2
L6	R	R	S	R	-	R	S	R	S	R	-	.89	.60 - 1.00	1	3	3	1
L7	R	R	S	S	-	S	S	S	S	R	-	.58	.39 - .78	1	2	6	1
L8	R	R	S	S	-	R	R	R	R	R	-	.41	.21 - .61	1	2	6	1
L9	R	R	S	S	-	S	R	S	S	R	-	.69	.18 - 1.00	1	1	1	1
L10	R	R	S	S	-	R	S	S	S	R	-	.79	.57 - 1.00	1	3	6	1
M1	R	R	S	S	-	R	R	R	-	R	-	.93	.85 - 1.00	5	10	26	5
M2	R	R	S	S	-	R	R	S	-	R	-	.99	.90 - 1.00	1	4	17	1
N1	R	R	S	S	-	R	S	R	-	R	S	.49	.15 - .84	1	1	2	1
N2	R	R	S	S	-	R	R	S	-	R	S	.82	.62 - 1.00	1	3	3	1
O1	R	R	S	S	-	R	R	S	S	R	S	.80	.70 - .90	3	8	18	3
O2	R	R	S	S	-	S	S	R	S	R	S	.76	.45 - 1.00	1	1	2	1
P	R	R	S	S	R	R	R	R	-	R	R	.94	.71 - 1.00	1	1	2	1
Q1	R	R	S	S	S	R	R	S	R	R	-	.71	.54 - .89	2	2	5	2
Q2	R	R	S	S	S	R	R	S	S	R	-	.99	.83 - 1.00	2	3	8	2
Q3	R	R	S	S	S	R	S	R	S	S	-	.99	.87 - 1.00	1	1	6	1
Q4	R	R	S	S	S	R	R	R	S	R	-	.90	.77 - 1.00	3	7	18	2

Table 8 (Continued)

Aggregated Results by Student Participation

Int. package marker	Components											Tau-U	95% CI	# of studies	# of subjects	# of ESs	# of unique geo-graphic regions
	Select DV	Def DV	Obs.	Rec.	Inst. pro.	Det. goal	Eval. goal	Sel. Rein	Admin sec. rein	Admin pri. rein	Graphing/ charting						
Q5	R	R	S	R	R	R	R	R	S	R	-	.97	.77 - 1.00	1	3	6	1
Q6	R	R	S	S	R	R	R	R	S	R	-	.05	-.26 - .36	1	2	4	1
R	R	R	S	S	S	R	S	S	R	R	S	.72	.46 - .99	1	1	2	1

Note. Select DV = Selection of dependent variable; Def DV = Definition of target behavior; Obs = Observation of target behavior; Rec. = Recording of behavior; Inst. pro. = Instructional prompts for target behavior; Det. goal = Determination of performance goal; Eval goal = Evaluation to determine whether performance goal was met; Admin sec. rein = Administration of secondary reinforcer; Admin pri rein = Administration of primary reinforcer; CI = Confidence interval. **These intervention sets meet evidence-based practice requirements. †These intervention sets approach evidence-based practice requirements, only lacking sufficient numbers of subjects.

In SM package L (Selection of the DV, Defining the DV, Observation, Recording, Selection of Reinforcers, Administration of Primary Reinforcers, Determining the Goal and Evaluating Goal Attainment), 10 variations of the intervention occurred within the literature. Four variations of SM package L (L2, L3, L6, & L9) could not be included in the current analysis because they had fewer than five ESs. Analysis was conducted on the six remaining intervention packages.

There was no significant difference (Kruskal-Wallis $p = .33$) between groups within SM package L. In SM package M (Selection of the DV, Defining the DV, Observation, Recording, Selection of Reinforcers, Administration of Primary Reinforcers, Determining the Goal and Evaluating Goal Attainment), two variations of the intervention occurred within the literature. A significant difference was detected between intervention packages (Wilcoxon $p = .004$). Within this analysis, SM packages that allowed the student to select the reinforcer (e.g., package M2) had improved effects over a similar intervention that allowed the researcher/teacher to select the reinforcer (e.g., package M1). In SM package N (Selection of the DV, Defining the DV, Observation, Recording, Selection of Reinforcers, Administration of Primary Reinforcers, Determining the Goal, Evaluating Goal Attainment, and Graphing Results), two variations of the intervention occurred within the literature. Differences between these intervention packages could not be evaluated because both versions of this intervention had fewer than five ESs. In SM package O (Selection of the DV, Defining the DV, Observation, Recording, Selection of Reinforcers, Administration of Primary Reinforcers, Administration of Secondary Reinforcers, Determining the Goal, Evaluating

Goal Attainment, and Graphing Results), two variations of the intervention occurred within the literature. However, SM package O2 could not be included in the current analysis because it had fewer than five ESs. Given the lack of sufficient data in package O2, no statistical significance testing was conducted within SM package O. In SM package Q (Selection of the DV, Defining the DV, Observation, Recording, Instructional Prompting, Selection of Reinforcers, Administration of Primary Reinforcers, Administration of Secondary Reinforcers, Determining the Goal, and Evaluating Goal Attainment), six variations of the intervention occurred within the literature. However, one variation of SM package Q (e.g. Q6) could not be included in the current analysis because it had fewer than five ESs. Analysis was conducted on the four remaining intervention packages. There was no significant difference (Kruskal-Wallis p = .89) between groups within SM package Q.

Research Question 4. Is SM Differentially Effective Based on Targeted Outcome?

The current analysis categorized five unique outcome variables within SM interventions (see Table 9). Within this analysis, Tau-U ESs ranged from a high of .95 CI_{95} [.89, 1.00] for Functional Communication to .64 CI_{95} [.60, .68] for Disruptive outcomes. The Kruskal-Wallis analysis showed significant differences between outcomes (p = <.0001). The Dunn post-hoc procedure (see Table 10) indicated significant differences between participants with the Functional Communication outcome and participants with the outcomes of On Task .80 CI_{95} [.78, .82], Disruptive .64 CI_{95} [.60, .68], and Social Behavior .73 CI_{95} [.68, .78]. Statistically significant differences were also found between participants with the Following Rules outcome and participants with On Task, Disruptive, Social Behavior outcomes.

Table 9

Aggregated Results by Outcome Variable

Dependent variable	Tau-U	95% CI	# of studies	# of subjects	# of analyses
On task	.80	.78 - .82	63	199	379
Disruptive	.64	.60 - .68	22	63	143
Social behavior	.73	.68 - .78	8	23	51
Functional communication	.95	.89 - 1.00	6	13	31
Following rules	.88	.80 - .95	6	27	31
Off task/disruptive	.73	.61 - .84	5	8	19

Table 10

Statistical Significance between Student Outcome Variables

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	Disruptive – Following rules	115.137	79.179	**
2	Disruptive – Functional communication	114.363	85.602	**
3	Disruptive – Off task/disruptive	25.345	119.107	
4	Disruptive – On task	33.853	46.338	
5	Disruptive – Social behaviors	14.086	65.921	
6	Following Rules – Functional communication	0.774	101.847	

Table 10 (Continued)

Statistical Significance between Student Outcome Variables

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
7	Following Rules – Off task/disruptive	89.793	131.269	
8	Following rules – On task	81.284	72.057	**
9	Following rules – Social behaviors	129.223	85.969	**
10	Functional communication – Off task/disruptive	89.018	135.24	
11	Functional communication – On task	80.51	79.061	**
12	Functional communication – Social behaviors	128.449	91.919	**
13	Off task/disruptive – On task	8.509	114.496	
14	Off task/disruptive – Social behaviors	39.43	123.725	
15	On task – Social behaviors	47.939	57.172	

Research Question 5. Is SM Differentially Effective Based on Participant Variables?

The current analysis examined the Age, Gender, and Disability categories to determine if differences occurred between studies on participant level variables (see Table 11). Within the Age variable, differences in ES magnitude were detected between participants at the Secondary .82 CI_{95} [.79, .85] and Primary .76 CI_{95} [.74, .78] levels. The differences were statistically significant (Wilcoxon p = .03). Within the Gender variable, small differences in ES magnitude were detected between Female .83 CI_{95} [.79, .87] and Male .76 CI_{95} [.74, .78] participants. These differences were not statistically significant (Wilcoxon p = .97).

Finally, within the Disability category variable, eight unique categories were identified in the literature. Within this analysis, Tau-U ESs ranged from .90 CI_{95} [.85, .95] for participants with Intellectual Disabilities to .66 CI_{95} [.61, .70] for participants classified as having an Emotional Disturbance. The Kruskal-Wallis analysis showed significant differences between participants of the Disability Category variable (p = <.0001). The Dunn post-hoc procedure (see Table 12) indicated significant differences between participants with Intellectual Disabilities and participants with Learning Disability .74 CI_{95} [.70, .78], At Risk .72 CI_{95} [.69, .75], or an Emotional Disturbance. Statistically significant differences were also found between the participants with ADHD .83 CI_{95} [.79, .88] and participants with a Learning Disability or considered At Risk. In addition, statistically significant difference were detected between students with Autism .87 CI_{95} [.83, .91] and participants labeled as At Risk.

Table 11

Aggregated Results by Student Characteristics

Student characteristic	Variable	Tau-U	95% CI	# of studies	# of subjects	# of analyses
Age/grade	Secondary	.82	.79 - .85	30	64	114
	Elementary	.76	.74 - .78	73	214	294
Gender	F	.83	.79 - .87	32	46	71
	M	.76	.74 - .78	84	225	329
Disability	ID	.90	.85 - .95	14	33	35
	AU	.87	.83 - .91	17	32	50
	TBI	.85	.62 - 1.00	1	3	3
	ADHD	.83	.79 - .88	12	35	60
	Gen Ed	.77	.65 - .88	5	12	12
	LD	.74	.70 - .78	24	61	71
	At risk	.72	.69 - .75	18	68	122
	ED	.66	.61 - .70	20	43	63

Note. Gender, F = Female, M = Male; Disability, ID = Intellectual Disability, AU = Autism spectrum disorder, TBI = Truamatic brain Iinjury, ADHD = Attention deficit hyperactivity disorder, LD = Learning disability, At risk = Children with behavior problems that are not formally classified, ED = Emotional/behavioral disturbance.

Table 12

Statistical Significance between Student Disability Categories

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	ADHD – At risk	76.2549	59.221	**
2	ADHD – AU	11.22	71.917	
3	ADHD – ED	66.7563	67.749	
4	ADHD – Gen ed	40.6333	118.768	
5	ADHD – LD	79.6697	65.861	**
6	ADHD – MR	18.5214	79.882	
7	ADHD – TBI	71.7167	222.194	
8	At risk – AU	65.0349	63.066	**
9	At risk – ED	9.4986	58.269	
10	At risk – Gen ed	35.6216	113.627	

Table 12 (Continued)

Statistical Significance between Student Disability Categories

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
11	At risk – LD	3.4148	56.062	
12	At risk – MR	94.7763	72.017	**
13	At risk – TBI	4.5383	219.489	
14	AU – ED	55.5363	71.135	
15	AU – Gen ed	29.4133	120.731	
16	AU – LD	68.4497	69.339	
17	AU – MR	29.7414	82.773	
18	AU – TBI	60.4967	223.25	
19	ED – Gen ed	26.123	118.295	
20	ED – LD	12.9134	65.006	
21	ED – MR	85.2778	79.178	**
22	ED – TBI	4.9603	221.942	
23	Gen ed – LD	39.0364	117.224	
24	Gen ed – MR	59.1548	125.638	
25	Gen ed – TBI	31.0833	242.433	
26	LD – MR	98.1911	77.569	**
27	LD –TBI	7.9531	221.373	
28	MR – TBI	90.2381	225.941	

Note. ID = Intellectually disability, AU = Autism spectrum disorder, TBI = Traumatic brain injury, ADHD = Attention deficit hyperactivity disorder, LD = Learning disability, At risk = Children with behavior problems that are not formally classified, ED = Emotional/behavioral disturbance.

Research Question 6. Is SM Differentially Effective Based on Student Instructional Setting?

The current analysis categorized three unique setting variables within SM interventions (see Table 13). Tau-U ESs ranged from of .87 CI_{95} [.79, .94] for Inclusion settings to .74 CI_{95} [.60, .68] for General Education settings. The Kruskal-Wallis analysis (see Table 14) showed no significant differences between settings (p = .25).

Table 13

Aggregated Results by Setting Variables

Setting Variable	Tau-U	95% CI	# of studies	# of subjects	# of analyses
Inclusion	.87	.79 - .94	9	19	23
Self-contained	.78	.75 - .81	42	102	123
Gen ed	.74	.72 - .77	44	157	223

*Note.*Gen ed = General Education.

Table 14

Statistical Significance between Setting Variables

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	Gen ed – Inclusion	37.1727	55.9236	
2	Gen ed – Self-contained	7.7194	28.6799	
3	Inclusion – Self-contained	29.4533	58.0101	

Research Question 7. Is SM Differentially Effective Based Cueing Strategies?

The current analysis categorized seven unique outcome variables within SM interventions (see Table 15). Tau-U ESs ranged from a high of 1.00 CI_{95} [.85, .1.00] for

Visual cues to .74 CI_{95} [.60, .68] for Audio cues. The Kruskal-Wallis analysis showed significant differences between outcomes (*p* = .0036). The Dunn post-hoc procedure (see Table 16) indicated significant differences between Peer cueing .99 CI_{95} [.84, 1.00] and Audio cueing .74 CI_{95} [.60, .68] methods.

Table 15

Aggregated Results by Cueing Variable

Cueing Variable	Tau-U	95% CI	# of studies	# of subjects	# of analyses
None	.79	.76 - .82	29	82	194
Audio	.74	.72 - .76	49	145	293
Teacher	.80	.75 - .85	10	36	118
Tactile	.89	.78 - 1.00	4	10	17
Visual	1.00	.85 - 1.00	1	3	12
Peer	.99	.84 - 1.00	2	7	9
Time	.78	.66 - .91	3	7	11

Table 16

Statistical Significance between Cueing Variables

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	Audio – None	20.615	40.807	
2	Audio – Peer	143.228	141.353	**
3	Audio – Tactile	50.243	119.194	
4	Audio – Teacher	40.559	54.602	
5	Audio – Time	30.293	119.194	
6	Audio – Visual	159.443	213.635	
7	None – Peer	122.613	142.034	
8	None – Tactile	29.628	120.001	
9	None – Teacher	19.944	56.34	

Table 16 (Continued)

Statistical Significance between Cueing Variables

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
10	None – Time	9.678	120.001	
11	None –Visual	138.828	214.086	
12	Peer – Tactile	92.986	180.875	
13	Peer – Teacher	102.669	146.594	
14	Peer – Time	112.936	180.875	
15	Peer – Visual	16.214	253.276	
16	Tactile – Teacher	9.683	125.365	
17	Tactile – Time	19.95	164.141	
18	Tactile – Visual	109.2	241.609	
19	Teacher – Time	10.267	125.365	
20	Teacher – Visual	118.883	217.139	
21	Time – Visual	129.15	241.609	

Research Question 8. Does Contingent Reinforcement of Behavior Improve Outcomes?

The current analysis examined differences between studies based on the use of contingent reinforcement (see Table 16). Within this variable, small differences in ES magnitude were detected between studies that used no reinforcement .78 CI_{95} [.76, .81] and studies that employed some form of contingent reinforcement .77 CI_{95} [.74, .79]. These differences were not statistically significant (Wilcoxon $p = .09$).

Table 17

Aggregated Results by Reinforcement Variable

Reinforcement Variable	Tau-U	95% CI	# of studies	# of subjects	# of analyses
None	.78	.76 - .81	48	155	288
Reinforcement for behavior	.77	.74 - .79	51	133	366

Research Question 9. Does Accuracy of Student Recording Relate to Improved Behavioral Outcomes?

This analysis categorized three unique variables within the SM Literature (see Table 17). Tau-U ESs ranged from of .83 CI_{95} [.81, .86] for studies with No Accuracy Check to .68 CI_{95} [.65, .71] for studies with an Accuracy Check. The Kruskal-Wallis analysis (see Table 18) showed no significant differences between outcomes based on checking participants' accuracy (p = .09).

Table 18

Aggregated Results by Accuracy Variable

Accuracy Variable	Tau-U	95% CI	# of studies	# of subjects	# of analyses
No accuracy check	.83	.81 - .86	62	182	316
Accuracy checked	.68	.65 - .71	13	45	128
Accuracy checked with reinforcement	.76	.73 - .79	22	63	210

Table 19

Statistical Significance between Accuracy Variables

Comparison number	Group comparisons	Difference in average ranks	Cutoff at alpha = 0.05	Significance difference = **
1	No accuracy check – Accuracy checked	25.5115	39.5649	
2	No Accuracy Check – Accuracy + reinforcement	25.0393	32.0278	
3	Accuracy checked – Accuracy + reinforcement	0.4722	42.6222	

CHAPTER V

DISCUSSION AND CONCLUSIONS

Teachers and students are in need of intervention protocols to help promote positive behavior that supports learning. SM is one such intervention that holds promise for building student capacity toward this goal. Despite the widespread use of this intervention and extensive research, several key procedural questions remain regarding use of SM interventions. For example, the number and type of key variables have varied across the research base on SM. In addition, the use of cueing, reinforcement, and accuracy of response have been similarly diverse. Accordingly, the goal of this meta-analysis was to examine factors related to implementation of SM protocols.

Self-Management as an EBP

This analysis generally confirmed the positive benefits of SM interventions and, as such, the results are consistent with previous research (Briesch & Chafouleas, 2009; Fantuzzo et al., 1987). However, considerable variability exists in the construction of SM interventions in the literature. Application of the Fantuzzo et al. (1987) theoretical framework showed 18 separate intervention component combinations under the heading of SM, and 44 combinations of components emerged based on student participation.

Not all SM intervention component constellations meet contemporary guidelines for classification as an EBP. EBP guidelines (Horner & Kratochwill, 2012) require that an intervention protocol must be based on a minimum of five studies that document adequate experimental control, these studies must be conducted by at least three research groups in three separate geographic locations, and must include effects for 20 separate participants.

Given the emphasis on replicable intervention procedures, the current analyses found that additional specificity is necessary to accurately describe the intervention procedure. Simply labeling interventions as SM does not accurately capture the variability that exists among researchers applying this intervention. The current analysis found that four combinations of intervention components separately met criteria as an EBP (see Table 4). These combined effects showed that among SM interventions that meet criteria for EBP, lower component SM interventions had similar or improved effects compared to SM packages with higher numbers of intervention components. This finding diverges from prior research, which showed no difference among SM intervention based on the total number of components (Briesch & Chafouleas, 2009; Fantuzzo et al., 1987).

Evaluation of Component Presence

The current analysis adds to the existing research base by evaluating studies with analogous intervention procedures. Previous analyses have only examined the total number of intervention components (Briesch & Chafouleas, 2009; Fantuzzo et al., 1987; Fantuzzo & Polite, 1990), which may have masked the finding that students show higher levels of improvement on SM interventions with fewer components.

Previous research had shown no differences between SM interventions based on the components used. In contrast, the current study found differences between SM interventions based on the components used. Specifically, two intervention packages, B

(Selection of the DV, Defining the DV, Observation, Recording, and Graphing and Charting) and I (Selection of the DV, Defining the DV, Observation, Recording, Instructional Prompting, Selecting Reinforcers, and Administering Primary and Secondary Reinforcers), showed significantly lower aggregate effects than many other intervention packages. The effects for SM package B are noteworthy because this package used only five components, a relatively small number in comparison to many of the other packages.

Despite this statistically significant result; however, effects for SM package B should be interpreted with caution because the effects for this package are based on a single study. In addition, it is doubtful that the addition of graphing and charting would have such deleterious effects in comparison to intervention packages without that component (e.g., SM package A). Given caution in interpreting the effects of SM package B, the results of the current analysis generally support the idea that in aggregate low-component SM packages are as effective as SM packages with more components.

The practical implication of this finding for school personnel is that SM interventions with fewer procedural steps are as effective as more complicated versions of the same intervention. Specifically, an SM intervention with four procedural steps (e.g., Selection of the DV, Defining the DV, Observation, and Recording) is as effective as other versions of this intervention with elaborate goal-setting and reinforcement procedures.

Student Involvement

An additional consideration in the evaluation of SM intervention components is student involvement. Similar to the variability in presence of intervention components, SM packages also vary widely in terms of student involvement. Within the research literature, 44 separate combinations of intervention components were identified when student involvement was applied to differentiate between components.

The current investigation found positive effects from studies that allowed students to self-record rather than having a teacher/researcher fulfill that role. These differences were significant within SM package D, and approached significance in SM package A. Although these results were not conclusive across all intervention packages, they support previous research (Briesch & Chafouleas, 2009) showing that student involvement in the recording of behavior supports more positive behavioral improvement.

With respect to the additional intervention components, the current study did not find any significant differences between student-directed and teacher/researcher-directed components. This may be partially explained by the finding that intervention components related to instructional prompting, determining/evaluating goals, selecting and applying reinforcement, and graphing are less essential in implementing SM interventions. Therefore, discriminating effects based on responsibility for implementing those components are less apparent due to the non-significant differences in implementing interventions with these components in general.

Behavioral Outcomes

Standards for EBP recommend direct evaluation of student outcomes to determine the degree to which an intervention improves socially valued outcomes. The current analysis found generally high effects across all outcome variables. Specifically, participants showed the highest effects on the Functional Communication and Following Rules outcome. Significant differences were detected between participants with the Functional Communication and other outcomes, such as On Task, Disruptive, and Social Behavior. Significant differences were also detected between Following Rules and Disruptive Behavior.

The difference between the Functional Communication and other outcomes is particularly noteworthy as functional communication tends to be emphasized for students with developmental disabilities (e.g., students with Intellectual Disability or Autism). This finding promotes SM intervention for a wide variety of outcomes and shows particular promise for promoting basic functional skills.

An additional noteworthy finding is the relatively smaller effects for Disruptive Behavior on aggregate. While participants made improvement in remediating Disruptive Behavior with this intervention, the effects on this outcome variable had a lower magnitude than other targeted outcomes. This may be because disruptive behavior is related to more intense deficits in self-control, a skill that is necessary for students to implement this intervention.

Student Characteristics

In terms of student outcomes, current EBP guidelines for SCR call for consideration of participant characteristics in evaluating intervention protocols (Horner & Kratochwill, 2012). These recommendations are based on the need to provide the most appropriate intervention for the individual. Previous analyses of these variables with respect to SM found no significant differences between participants based on Gender, Age, and Disability Status (Briesch & Chafouleas, 2009).

The current analysis also examined Gender, Age, and Disability Status and found significant differences within the Age and Disability Status variables. No significant differences were detected based on student Gender. With regard to Age, improved intervention effects were noted for secondary-level students compared to primary-level students. Although high effects were found in aggregate at the primary level, the aggregate effects at the secondary level are noteworthy. The increased effects at the secondary level may be influenced by the higher level of developmental maturity in these students, which may enable them to benefit more from the self-focused tasks involved in this intervention.

The current analysis also found noteworthy differences based on Disability Status. In terms of magnitude, the highest effects were found for students with ID, AU, TBI, and ADHD. SM interventions appear well suited for students with developmental disabilities such as ID and AU, possibly due a better match between the intervention methods and the strengths of students with these disability categories. Although this may not be the case for all students with developmental disabilities, students with ID or AU, on aggregate, seemed to respond more positively to the increased prompting and performance feedback available in this intervention. By creating a structure for students with developmental disabilities to prompt positive behavior and evaluate themselves, the current study found notably positive improvements for students with these disability profiles.

In contrast, relatively lower magnitude effects were found for participants with No Disability, LD, ED and those considered At Risk. Of particular interest are the relatively lower effects for participants identified as At Risk and ED. While the current findings are not conclusive, there is some indication that these self-focused interventions may not be as effective for students with these disability profiles. These results support previous findings that students with ED have generally poor long-term outcomes despite extensive intervention in schools (Bradley, Doolittle, & Bartolotta, 2008; Wagner et al., 2006).

Instructional Setting

Instructional setting is an important consideration for educators seeking to apply behavioral interventions. Thus, the instructional setting has the potential to impact student to-teacher ratios and teachers' capacity to attend to intervention procedures for individual students. While SM interventions have been effectively applied across a variety of instructional settings over the years, this variable has not been examined for the purpose of determining if specific differences exist in student outcomes based on the instructional setting.

The current study found no differences between behavioral outcomes based on instructional setting, thus supporting the use of SM across settings. Federal education code (IDEIA, 2004; NCLB, 2001) mandates that students be educated in the least restrictive environment. The current analysis showed that SM interventions are effective in general education as well as inclusive settings. Considering the positive benefits of educating children in general education settings due to increased access to grade-appropriate educational material, SM holds promise as a behavioral intervention that integrates well into these settings.

Methods of Cueing

An additional aspect of SM that has been examined and manipulated in previous research is the method of cueing used. Recommendations regarding the use of cueing have differed in previous research. Some have noted that cueing is an important procedure that promotes positive outcomes with SM interventions (Hallahan & Sapona, 1983; Heins et al., 1986). Others have questioned the utility of externally controlled methods in promoting self-regulatory behavior (Snider, 1987).

The current analysis did not find a significant difference between studies that used no form of cueing and studies that employed some form of cueing. Thus, it does not settle the debate of whether cueing is necessary for student reactivity to SM interventions, as one method was not found to be superior to another. This nonsignificant difference confirms that implementers of SM have some flexibility in applying cueing methods.

This finding holds important practical implications for the use of SM interventions in the classroom. In terms of teacher involvement, no differences were found between studies that required the teacher to cue the participant and other forms of cueing. Therefore, teacher time and involvement in cueing or managing the cueing schedule may not improve student outcomes over other forms of cueing or no cueing at all.

Within the analysis of cueing methods, one significant result was detected between Peer cueing and Audio cueing. This finding should be interpreted with caution, however, since only two studies employed the Peer cueing method. Nevertheless, this finding shows a potentially positive impact of peer involvement in SM interventions.

Contingent Reinforcement

The current analysis also examined the value of contingent reinforcement in promoting positive behavior change. Contingent reinforcement is used ubiquitously in education to promote positive outcomes for children (Cameron & Pierce, 1994). Indeed, in the current study, 12 of the 18 package combinations included some form of contingent reinforcement. Given the known positive effects that accompany the use of contingent reinforcement, the current application examined whether these positive effects translate to positive additive effects for SM interventions.

Virtually no difference was found between studies based on the use of contingent reinforcement. This finding appears fairly robust given the number of studies examined with both the No Reinforcement (48) and Contingent Reinforcement (51) condition. Thus, it supports previous assertions (Snider, 1987) that contingent reinforcement is not necessary to promote positive outcomes with SM interventions.

Accuracy in Recording

Student accuracy in recording has also been debated in terms of its contribution to student outcomes using SM interventions. Intuitively, one would tie student behavior improvement to accurate recording of behavior. Thus, accuracy of recording has long accompanied this intervention as a strategy to ensure that student recording of behavior matches the outcome (Blick & Test, 1987; Rumsey & Ballard, 1985; Smith et al., 1988).

However, despite recommendations for this practice within the literature, many have maintained that accuracy in student recording is not a necessary feature of this intervention (Nelson & Hayes, 1981; Rooney, Hallahan, & Lloyd, 1984). The current analysis supports the position that accuracy of recording does not enhance behavior outcomes. In fact, there was no significant difference between studies that did not include accuracy check versus studies with accuracy checks or studies with accuracy checks that were tied to reinforcement. This finding supports previous conclusions that participant reactivity to SM is not based on external pressure to accurately recording of behavior. Instead, as posited by Nelson and Hayes (1981), it may be the SM intervention that “serves to cue ultimate environmental consequences” (p. 9). That is, it may be the environmental consequences that produce behavioral change rather than an imposition of accurate recording.

Limitations

The current analysis is affected by some limitations that should be considered when interpreting results. First, the study relied on post-hoc analyses of published SCR study data to calculate the ES. The field of SCR has not yet reached consensus regarding the most appropriate method to calculate an effect size. While meta-analyses have been published using the Tau-U effect size (Bowman-Perrot et al., 2013), the current study represents the largest application of the Tau-U effect size to date.

Several methods are being promoted for the purpose of analyzing SCR data. Primary among these methods are randomization methods (Edgington, 1975; Kratochwill & Levin, 2010), hierarchical models (Beretvas & Chung 2008; Van den Noortgate & Onghena, 2008), and nonoverlap models (Parker, Vannest, & Davis, 2011). Nonoverlap models, as used in the current analysis, have been criticized (Wolery, Busick, Reichow, & Barton, 2010) for not adequately expressing the magnitude of change. While this criticism has some merit, as the Tau-U method has a limited range of scores (i.e., -1 to 1), the limited range of values for this statistic is a function of the use of proportions as the unit of analysis. It must be acknowledged that the use of proportions does limit the expression of effects beyond 100% improvement, but this unit of analysis allows a greater degree of flexibility in application to SCR.

The use proportions in the calculation of an effect sizes eliminates the influence of the scale of measurement (e.g., y-axis scale). Many SCR studies differ in how the dependent variable is measured. For example, one study may employ a Likert-type rating scale to evaluate On Task behavior; whereas, another may measure the same outcome by counting the percent of intervals. In order to compare and aggregate studies with differing scales of measurement, a conversion must occur. Nonoverlap methods (e.g., Tau-U) accomplish this goal by using proportions as the unit of analysis. While this method does limit the range of effect size expression, there is value added through this conversion in the ability to compare and aggregate effect sizes across studies with differing scales of measurement.

An additional limitation of the current study is the classification of student-level variables used. Within meta-analysis, researchers rely on detailed and accurate reporting of variables of interest within the text of the study. Significant differences were apparent between studies in how disability categories were conceptualized and labeled. Also, differences between clinical and educational classification taxonomy cause a lack of continuity in the classification of students.

This issue was most evident in the classification of the At Risk disability category. Several students with fairly severe clinical diagnoses were classified as being At Risk because an educational classification of ED was not present or reported. Given the similarities in effect between the At Risk and ED disability categories, the results of the current analysis may reflect the lack of specificity inherent in the current educational classification system

A further limitation of the study is due to the breadth of analyses. The research sought to categorize studies based on the methods used to apply the intervention. However, the analysis revealed that a great deal of variability exists between studies based on the components used to apply a SM intervention.

This issue was most evident on in the application of student participation in intervention components. Thus, 44 different combinations of intervention components were found when student participation was considered. Given this large number of variables and lack of continuity between studies, differences between studies could not be reasonably detected. This is primarily due to the degree of correction that is necessary when considering experiment-wise error.

The final limitation of the study relates to potential interaction effects among variables. Only the identified variables were analyzed. The interactions among them were not considered. Interaction between variables may provide additional information to promote practices for certain students. For example, the aggregate effect of reinforcement was nonsignificant; however, this effect may not be true for certain disability categories.

Future analyses should examine the differential effects of each of the moderator variable on each of the student characteristic variables to determine if differences exist between SM intervention methods based on specific student characteristics.

Conclusions

The current study examined the literature related to SM interventions in an attempt to provide guidance for educators considering this intervention. This analysis illuminated several important factors to be considered when conceptualizing and applying SM interventions in schools.

Of primary importance, the study found that SM interventions meet current SCR criteria for classification as an EBP. While all possible iterations of this intervention do not contain sufficient evidence to justify this classification, four separate versions of the SM intervention meet all criteria for classification as an EBP. Within these four versions, protocols with fewer components outperformed protocols with more components in aggregate. In addition, protocols with higher levels of student participation showed higher effects for improving behavior.

Analyses across all possible iterations of the SM intervention confirm the finding that the number of components applied in SM interventions does not seem to impact overall outcome. Thus, the current study found that a basic 4-component SM intervention had similar effects to a more complex 11-component intervention packages.

With regard to student-level factors, the current analysis found that SM interventions showed higher effects, on aggregate, in secondary-level students. In addition, students with developmental disabilities (e.g., ID and AU) demonstrated greater improvements than other student disability categorizations. Further, evaluation of school settings showed that SM interventions had equivalent effects across a variety of instructional arrangements.

Analysis of specific moderator variables demonstrated that the use of contingent reinforcement does not provide additive effects to this intervention. In addition, no significant differences were detected between various methods to cue student recording. Moreover, studies with no cuing of student recording had similar effects to studies that used teacher dependent cuing methods. Accuracy in student recording was found to be nonessential in promoting behavioral change. As such, active external monitoring of the student recording product may not function to improve behavioral outcomes.

The goal of SM interventions is to maximize student functioning in classroom settings for children from a variety of disability categories. The current meta-analysis provides evidence that SM interventions meet criteria as an evidence-based practice (Horner & Kratochwill, 2012) to achieve this goal. To this end, the study provides evidence that more streamlined (i.e. lower component) versions of SM interventions have similar or improved effects compared to more complicated versions of the same intervention. Although each student's needs should be carefully considered when implementing any behavioral intervention, the current analysis shows that, on aggregate, a SM intervention with four basic components is as effective as SM interventions that include additional components.

REFERENCES

Agran, M., Sinclair, T., Alper, S., Cavin, M., Wehmeyer, M., & Hughes, C. (2005). Using self-monitoring to increase following-direction skills of students with moderate to severe disabilities in general education. *Education and Training in Developmental Disabilities*, *40*(1), 3-13.

Agresti, A. (2010). *Analysis of ordinal categorical data* (2nd ed.). Hoboken, NJ: Wiley.

Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy*, *31*(6), 621-631.

Altman, D. G. (1991). *Practical statistics for medical research*. London, UK: Chapman & Hall/CRC.

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools*, *43*, 211-221.

Ary, D., & Suen, H. K. (1989). *Analyzing qualitative behavioral observational data*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91-97.

Bandura, A. (1969). *Principles of behavior modification*. New York, NY: Holt, Rinehart & Winston

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Barling, J., & Patz, M. (1980). Differences following self-and external-reinforcement as a function of locus of control and age: A social learning analysis. *Personality and Individual Differences, 1*(1), 79-85.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173-1182.

Barry, L. M., & Messer, J. J. (2003). A practical application of self-management for students diagnosed with attention-deficit/hyperactivity disorder. *Journal of Positive Behavior Interventions, 5,* 238-248.

Belfiore, P. J., Fritts, K. M., & Herman, B. C. (2008). The role of procedural integrity using self-monitoring to enhance discrete trial instruction (DTI). *Focus on Autism and Other Developmental Disabilities*, *23*(2), 95-102.

Beretvas, S. N., & Chung, H. (2008). A review of meta-analyses of single-subject experimental designs: Methodological issues and practice. *Evidence-Based Communication Assessment and Intervention*, *2*(3), 129-141.

Blick, D. W., & Test, D. W. (1987). Effects of self-recording on high-school students' on-task behavior. *Learning Disability Quarterly*, 203-213.

Bornstein, P. H., & Quevillon, R. P. (1976). The effects of a self-instructional package on overactive preschool boys. *Journal of Applied Behavior Analysis*, *9*(2), 179-188.

Bowers, D. S., Clement, P. W., Fantuzzo, J. W., & Sorensen, D. A. (1985). Effects of teacher-administered and self-administered reinforcers on learning disabled children. *Behavior Therapy*, *16*(4), 357-369.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review, 42*(1), 39-55.

Boyle, J. R., & Hughes, C. A. (1994). Effects of self-monitoring and subsequent fading of external prompts on the on-task behavior and task productivity of elementary students with moderate mental retardation. *Journal of Behavioral Education*, *4*(4), 439-457.

Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education*, *17*(1), 4-23.

Briesch, A. M., & Chafouleas, S. M. (2009). Review and analysis of literature on self-management interventions to promote appropriate classroom behaviors (1988-2008). *School Psychology Quarterly, 24*, 106-118.

Broden, M., Hall, R. V., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eighth-grade students. *Journal of Applied Behavior Analysis, 4,* 191-199.

Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions*, *5*(3), 144-152.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived selfefficacy in classroom management. *Teaching and Teacher Education*, *16*(2), 239-254.

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., & Erkanli A. (1996). Children's mental health service use across service sectors. *Health Affairs, 14*(3), 147-159.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64,* 363-423.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification*, *28*(2), 234-246.

Cavalier, A. R., Ferretti, R. P., & Hodges, A. E. (1997). Self-management within a classroom token economy for students with learning disabilities. *Research in Developmental Disabilities, 18,* 167-178.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685-716

Christie, D. J., Hiss, M., & Lozanoff, B. (1984). Modification of inattentive classroom behavior: Hyperactive children's use of self-recording with teacher guidance. *Behavior Modification*, *8*(3), 391-406.

Clees, T. J. (1994). Self-recording of students' daily schedules of teachers' expectancies: Perspectives on reactivity, stimulus control, and generalization. *Exceptionality*, *5*(3), 113-129.

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin, 114*, 494-509.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, R., Rubin, S., & Heinen, J. R. (1979). Collateral increases in performance of long division resulting from on-task gains using token reinforcement and self-recording. *Psychological Reports*, *44*(2), 651-658.

Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *School Psychology Quarterly*, *22*(4), 540-556.

Crawley, S. H., Lynch, P., & Vannest, K. (2006). The use of self-monitoring to reduce off-task behavior and cross-correlation examination of weekends and absences as an antecedent to off-task behavior. *Child & Family Behavior Therapy*, *28*(2), 29-48.

Dalton, T., Martella, R. C., & Marchand-Martella, N. E. (1999). The effects of a self-management program in reducing off-task behavior. *Journal of Behavioral Education*, *9*(3-4), 157-176.

Delaney, H. D., & Vargha, A. (2002). Comparing two groups by simple graphs. *Psychological Bulletin, 79*, 110-116.

DiGangi, S. A., Maag, J. W., & Rutherford, R. B. (1991). Self-graphing of on-task behavior: Enhancing the reactive effects of self-monitoring on on-task behavior and academic performance. *Learning Disability Quarterly, 14,* 221-230.

Dunlap, G., Clarke, S., Jackson, M., Wright, S., Ramos, E., & Brinson, J. (1995). Self-monitoring of classroom behaviors with students exhibiting emotional and behavioral challenges. *School Psychology Quarterly, 10,* 165-177.

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*,*6*(3), 241-252.

Edgington, E. S. (1975). Randomization tests for one-subject operant experiments. *The Journal of Psychology*, *90*(1), 57-68.

Fantuzzo, J. W., & Polite, K. (1990). School-based, behavioral self-management: A review and analysis. *School Psychology Quarterly, 5,* 180-198.

Fantuzzo, J. W., Rohrbeck, C. A., & Azar, S. T. (1987). A component analysis of behavioral self management interventions with elementary school students. *Child & Family Behavior Therapy, 9,* 33-43.

Franz, D. P., Vannest, K. J., Parker, R. I., Hasbrouck, J. E., Dyer, N., & Davis, J. L. (2008). Time use by special education teachers and how it is valued. *Journal of School Leadership. 18*(5), 551-576.

Freeman, K. A., & Dexter-Mazza, E. T. (2004). Using self-monitoring with an adolescent with disruptive classroom behavior. *Behavior Modification*, *28*, 402-419.

Ganz, J. B., & Sigafoos, J. (2005). Self-monitoring: Are young adults with MR and autism able to utilize cognitive strategies independently? *Education and Training in Developmental Disabilities*, *40*, 24-33.

GetData. (2012). *Get data graph digitizer* [Computer software]. Retrieved from http://getdata-graph-digitizer.com/index.php

Gillam, S. L., & Gillam, R. B. (2006). Making evidence-based decisions about child language intervention in schools. *Language, Speech, and Hearing Services in Schools*, *37*(4), 304-315.

Guresko-Moore, S., DuPaul, G. J., & White, G. P. (2007). Self-management of classroom preparedness and homework: Effects on school functioning of adolescents with attention deficit hyperactivity disorder. *School Psychology Review*, *36*, 647-664.

Hallahan, D. P., & Sapona, R. (1983). Self-monitoring of attention with learning disabled children: Past research and current issues. *Journal of Learning Disabilities*, *15*, 616-620.

Hallahan, D. P., Lloyd, J. W., Kneedler, R. D., & Marshall, K. J. (1982). A comparison of the effects of self versus teacher assessment of on-task behavior. *Behavior Therapy*, *13*, 715-723.

Harrison, J. R., Vannest, K., Davis, J., & Reynolds, C. (2012). Common problem behaviors of children and adolescents in general education classrooms in the United States. *Journal of Emotional and Behavioral Disorders*, *20*(1), 55-64.

Heins, E. D., Lloyd, J. W., & Hallahan, D. P. (1986). Cued and non-cued self-recording of attention to task. *Behavior Modification*, *10*(2), 235-254.

Hoff, K. E., & DuPaul, G. J. (1998). Reducing disruptive behavior in general education classrooms: The use of self-management strategies. *School Psychology Review*, *27*, 290-303.

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods*. New York, NY John Wiley & Sons.

Holman, J., & Baer, D. M. (1979). Facilitating generalization of on-task behavior through self-monitoring of academic tasks. *Journal of Autism and Developmental Disorders*, *9*(4), 429-446.

Horner, R. H., & Kratochwill, T. R. (2012). Synthesizing single-case research to identify evidence-based practices: Some brief reflections. *Journal of Behavioral Education*, *21*(3), 266-272.

Horner, R. H., Carr, E. G., Halle, J., McGee, J., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(2), 165-179.

Hughes, C., Copeland, S. R., Agran, M., Wehmeyer, M. L., Rodi, M. S., & Presley, J. A. (2002). Using self-monitoring to improve performance in general education high school classes. *Education and Training in Mental Retardation and Developmental Disabilities*, *37*(3), 262-272

Hughes, C., & Hendrickson, J. (1987). Self-monitoring with at-risk students in the regular class setting. *Education and Treatment of Children, 10*, 225-236.

Hughes, C. A., Ruhl, K. L., & Misra, A. (1989). Self-management with behaviorally disordered students in school settings: A promise unfulfilled? *Behavioral Disorders, 14,* 250-262.

Individuals With Disabilities Education Act of 1997, Pub. L. No. 105–17 [Amending 20 U.S.C. § 1400 et seq.].

Individuals With Disabilities Education Improvement Act of 2004, Pub. L. No. 108– 446, 118 Stat. 2647 [Amending 20 U.S.C. § 1462 et seq.].

Kanfer, F. H. (1977). The many faces of self-control, or behavior modification changes its focus. In R. B. Stuart (Ed.), *Behavior self-management: Strategies, techniques, and outcomes* (pp. l-48). New York, NY: Brunner/Mazel.

Kataoka, S., Rowan, B., & Hoagwood, K. E. (2009). Bridging the divide: In search of common ground in mental health and education research and policy. *Psychiatric Services*, *60*(11), 1510-1515.

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods*, *15*(2), 124-144.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation.* Retrieved from: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583-621.

Lagomarcino, T. R., & Rusch, F. R. (1989). Utilizing self-management procedures to teach independent performance. *Education and Training in Mental Retardation*, *24*(4), 297-305.

Landrum, T. J., Tankersley, M., & Kaufman, J. M. (2003). What is special about special education for students with emotional or behavioral disorders? *The Journal of Special Education, 37,* 148-156.

Lee, S. H., Simpson, R. L., & Shogren K. A. (2007). Effects and implications of self-management for students with autism: A meta-analysis. *Focus on Autism and Other Developmental Disabilities*, 22, 2-13. doi: 10.1177/10883576070220010101

Levendoski, L. S., & Cartledge, G. (2000). Self-monitoring for elementary school children with serious emotional disturbances: Classroom applications for increased academic responding. *Behavioral Disorders*, *25,* 211-224.

Lipinski, D., & Nelson, R. (1974). The reactivity and unreliability of self-recording. *Journal of Consulting and Clinical Psychology, 42,* 118-123.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lloyd, J. W., Hallahan, D. P., Kosiewicz, M. M., & Kneedler, R. D. (1982). Reactive effects of self-assessment and self-recording on attention to task and academic productivity. *Learning Disability Quarterly*, 216-227.

Lloyd, M. E., & Hilliard, A. M. (1989). Accuracy of self-recording as a function of repeated experience with different self-control contingencies. *Child & Family Behavior Therapy, 11(2),* 1-14.

Mace, F. C., & Kratochwill, T. R. (1985). Theories of reactivity in self-monitoring. *Behavior Modification, 9*, 323-343.

Mace, F. C., Belfiore, P. J., & Hutchinson, J. M. (2001). Operant theory and research on self-regulation. In B. Zimmerman & D. Schunk (Eds.), *Self regulated learning and academic achievement* (pp. 39-66). New York, NY: Guilford.

Mace, F. C., Shapiro, E. S., West, B. J., Campbell, C., & Altman, J. (1986). The role of reinforcement in reactive self-monitoring. *Applied Research in Mental Retardation, 7*, 315-327.

Marquis, J. G., Horner, R. H., Carr, E. G., Turnbull, A. P., Thompson, M., & Behrens, G. A. (2000). A meta-analysis of positive behavior support. In R. Gersten, E. Schiller, & S. Vaughn, (Eds.), *Contemporary special education research: Syntheses of knowledge base on critical instructional issues*. (pp. 137-178) Mahwah, NJ: Erlbaum.

McDougall, D., & Brady, M. P. (1995). Using audio-cued self-monitoring for students with severe behavior disorders. *Journal of Educational Research, 88*, 309-318.

Meichenbaum, D. H., & Goodman, J. (1971). Training impulsive children to talk to themselves: a means of developing self-control. *Journal of Abnormal Psychology*, *77*(2), 115-126.

Mooney, P., Ryan, J. B., Uhing, B. M., Reid, R., & Epstein, M. H. (2005). A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education, 14,* 203-221.

Morris, S., & Messer, S. B. (1978). The effect of locus of control and locus of reinforcement on academic task persistence. *The Journal of Genetic Psychology*, *132*(1), 3-9.

Morrow, L. W., Burke, J. G., & Buell, B. J. (1985). Effects of a self-recording procedure on the attending to task behavior and academic productivity of adolescents with multiple handicaps. *Mental Retardation*, *23*(3), 137-141.

Nelson, R. O., & Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification, 5*(1), 3-14.

Nelson, R. O., Lipinski, D. P., & Black, J. L. (1975). The relative reactivity of external observations and self-monitoring. *Behavior Therapy, 7*, 314-321.

Newman, B., Buffington, D. M., & Hemmes, N. S. (1996). Self-reinforcement used to increase the appropriate conversation of autistic teenagers. *Education and Training in Mental Retardation and Developmental Disabilities, 31*, 304-309.

Newman, B., Reinecke, D. R., & Meinberg, D. L. (2000). Self-management of varied responding in three students with autism. *Behavioral Interventions, 15,* 145-151.

Ninness, H. A., Fuerst, J., Rutherford, R. D., & Glenn, S. S. (1991). Effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. *Journal of Applied Behavior Analysis*, *24*(3), 499-508.

No Child Left Behind (NCLB) Act of 2001, 20 U.S.C.A. § 6301.

Parker, R. I. (2006). Increased reliability for single case research results: Is the bootstrap an answer? *Behavior Therapy*, 37(4), 326-338.

Parker, R., & Hagan-Burke, S. (2007). Single case research results as clinical Outcomes. The *Journal of School Psychology*, *45,* 637-653.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single case research. *Exceptional Children*, *75*, 135-150.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, *35*(4), 303-322.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284-299.

Parker, R. I., Brossart, D. F., Callicott, K. J., Long, J. R., De-Alba, R. G., Baugh, F. G., & Sullivan, J. R. (2005). Effect sizes in single case research: How large is large? *School Psychology Review, 34*(1), 116-132.

Rachlin, H. (1974). Self-control. *Behaviorism, 2,* 94-107.

Rasing, E. J., Coninx, F. C., Duker, P. C., & van den Hurk, A. J. (1994). Acquisition and generalization of social behaviors in language disabled deaf adolescents. *Behavior Modification*, *18*(4), 411-442.

Reid, R. (1996) Research in self-monitoring with students with learning disabilities: The present, the prospect, the pitfalls. *Journal of Learning Disabilities, 29*(3), 317-331.

Reinecke, D. R., Newman, B., & Meinberg, D. L. (1999). Self-management of sharing in three pre-schoolers with autism. *Education and Training in Mental Retardation and Developmental Disabilities*, *34*, 312–317.

Rooney, K. J., & Hallahan, D. P. (1988). The effects of self-monitoring on adult behavior and student independence. *Learning Disabilities Research, 3,* 88-93.

Rooney, K. J., Hallahan, D. P., & Lloyd, J. W. (1984). Self-recording of attention by learning disabled students in the regular classroom. *Journal of Learning Disabilities*, *17*(6), 360-364.

Rumsey, I., & Ballard, K. D. (1985). Teaching self-management strategies for independent story writing to children with classroom behaviour difficulties. *Educational Psychology*, *5*(2), 147-157.

Sainato, D. M., Goldstein, H., & Strain, P. S. (1992). Effects of self-evaluation on preschool children's use of social interaction strategies with their classmates with autism. *Journal of Applied Behavior Analysis, 25*(1), 127-141.

Shearer, D. D., Kohler, F. W., Buchan, K. A., & McCullough, K. M. (1996). Promoting independent interactions between preschoolers with autism and their nondisabled peers: An analysis of self-monitoring. *Early Education and Development, 7,* 205-220.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioural sciences.* New York, NY: McGraw-Hill.

Smith, B. W., & Sugai, G. (2000). A self-management functional assessment-based behavior support plan for a middle school student with EBD. *Journal of Positive Behavior Interventions, 2,* 208-217.

Smith, D. J., Young, K. R., West, R. P., Morgan, D. P., & Rhode, G. (1988). Reducing the disruptive behavior of junior high school students: A classroom self-management procedure. *Behavioral Disorders, 13,* 231-239.

Snider, V. (1987). Use of self-monitoring of attention with LD students: Research and application. *Learning Disability Quarterly, 10,* 139-151.

Strain, P. S., Kohler, F. W., Storey, K., & Danko, C. D. (1994). Teaching preschoolers with autism to self-monitor their social interactions: An analysis of results in home and school settings. *Journal of Emotional and Behavioral Disorders, 2*, 78-88.

Tate, R.L., Mcdonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials:

Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation*, *18*(4), 385-401.

United States Census Bureau. (2012). *Geography reference.* Retrieved from http://www.census.gov/geo/reference/.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention, 2*(3), 142-151.

Vannest, K. J., Reynolds, C., & Kamphaus, R. (2008). *BASC-2 intervention guide for learning and behavior problems*. Minneapolis, MN: Pearson.

Vannest, K. J., Mahadevan, L., Mason, B. A Temple-Harvey, K.., &. (2009). Educator perceptions of the impact of No Child Left Behind on special populations. *Remedial and Special Education, 30*(3), 148-159.

Wagner, M., Friend, M., Bursuck, W. D., Kutash, K., Duchnowski, A. J., Sumi, W. C., & Epstein, M. H. (2006). Educating students with emotional disturbances: A national perspective on school programs and services. *Journal of Emotional and Behavioral Disorders, 14*(1), 12-30.

Webber, J., Schuermann, B., McCall, C., & Coleman, M. (1993). Research on self-monitoring as a behavior management technique in special education classrooms: A descriptive review. *Remedial and Special Education, 14*(2), 38-56.

Wilcox, R. R. (2010). *Fundamentals of modern statistical methods: Substantially improving power and accuracy*. New York, NY: Springer.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80-83.

Wolery, M., Busick, M., Reichow, B., & Barton, E. E. (2010). Comparison of overlap methods for quantitatively synthesizing single-subject data. *The Journal of Special Education*, *44*(1), 18-28.

Wood, S. J., Murdock, J. Y., & Cronin, M. E. (2002). Self-monitoring and at-risk middle school students: Academic performance improves, maintains, and generalizes. *Behavior Modification, 26,* 605-626.

Zegiob, L., Klukas, N., & Junginger, J. (1978). Reactivity of self-monitoring procedures with retarded adolescents. *American Journal of Mental Deficiency*, *83*(2), 156-163.

APPENDIX A

STUDIES INCLUDED IN ANALYSIS

Agran, M., Sinclair, T., Alper, S., Cavin, M., Wehmeyer, M., & Hughes, C. (2005). Using self-monitoring to increase following-direction skills of students with moderate to severe disabilities in general education. *Education and Training in Developmental Disabilities*, *40*(1), 3-13.

Alberto, P. A., Taber, T. A., & Fredrick, L. D. (1999). Use of self-operated auditory prompts to decrease aberrant behaviors in students with moderate mental retardation. *Research in developmental disabilities*, *20*(6), 429-439.

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools*, *43*(2), 211-221.

Barry, L. M., & Messer, J. J. (2003). A practical application of self-management for students diagnosed with attention-deficit/hyperactivity disorder. *Journal of Positive Behavior Interventions*, *5*(4), 238-248.

Bowers, D. S., Clement, P. W., Fantuzzo, J. W., & Sorensen, D. A. (1985). Effects of teacher-administered and self-administered reinforcers on learning disabled children. *Behavior Therapy, 16*(4), 357-369.

Boyle, J. R., & Hughes, C. A. (1994). Effects of self-monitoring and subsequent fading of external prompts on the on-task behavior and task productivity of elementary students with moderate mental retardation. *Journal of Behavioral Education*, *4*(4), 439-457.

Broden, M., Hall, R. V., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eigth-grade students. *Journal of Applied Behavior Analysis*, *4*(3), 191-199.

Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions*,*5*(3), 144-152.

Callahan, K., & Rademacher, J. A. (1999). Using self-management strategies to increase the on-task behavior of a student with autism. *Journal of Positive Behavior Interventions*, *1*(2), 117-122.

Christensen, L., Young, K. R., & Marchant, M. (2004). The Effects of a Peer-Mediated Positive Behavior Support Program on Socially Appropriate Classroom Behavior. *Education & Treatment of Children*, *27*(3), 199-234.

Christie, D. J., Hiss, M., & Lozanoff, B. (1984). Modification of Inattentive Classroom Behavior Hyperactive Children's Use of Self-Recording with Teacher Guidance. *Behavior Modification*, *8*(3), 391-406.

Cihak, D. F., Wright, R., & Ayres, K. M. (2010). Use of self-modeling static-picture prompts via a handheld computer to facilitate self-monitoring in the general education classroom. *Education and Training in Developmental Disabilities*, *45*(1), 136-149.

Clees, T. J. (1994). Self-recording of students' daily schedules of teachers' expectancies: Perspectives on reactivity, stimulus control, and generalization. *Exceptionality*, *5*(3), 113-129.

Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group Contingencies, Randomization of Reinforcers, and Criteria for Reinforcement, Self-Monitoring, and Peer Feedback on Reducing Inappropriate Classroom Behavior. *School Psychology Quarterly*, *22*(4), 540-556.

Coyle, C., & Cole, P. (2004). A videotaped self-modelling and self-monitoring treatment program to decrease off-task behaviour in children with autism. *Journal of Intellectual and Developmental disability*, *29*(1), 3-16.

Crawley, S. H., Lynch, P., & Vannest, K. (2006). The use of self-monitoring to reduce off-task behavior and cross-correlation examination of weekends and absences as an antecedent to off-task behavior. *Child & family behavior therapy*, *28*(2), 29-48.

Dalton, T., Martella, R. C., & Marchand-Martella, N. E. (1999). The effects of a self-management program in reducing off-task behavior. *Journal of Behavioral Education*, *9*(3-4), 157-176.

Davies, S. C., Jones, K. M., & Rafoth, M. A. (2010). Effects of a self-monitoring intervention on children with traumatic brain injury. *Journal of Applied School Psychology*, *26*(4), 308-326.

Davies, S., & Witte, R. (2000). Self-management and peer-monitoring within a group contingency to decrease uncontrolled verbalizations of children with Attention-Deficit/hyperactivity disorder. *Psychology in the Schools*, *37*(2), 135-147.

DiGangi, S. A., & Maag, J. W. (1992). A Component Analysis of Self-Management Training with Behaviorally Disordered Youth. *Behavioral Disorders*,*17*(4), 281-290.

DiGangi, S. A., Maag, J. W., & Rutherford Jr, R. B. (1991). Self-graphing of on-task behavior: Enhancing the reactive effects of self-monitoring on on-task behavior and academic performance. *Learning Disability Quarterly*, 221-230.

Dunlap, G. (1995). Self-monitoring of classroom behaviors with students exhibiting emotional and behavioral challenges. *School Psychology Quarterly*,*10*(2), 165-177.

Edwards, L., Salant, V., Howard, V. F., Brougher, J., & McLaughlin, T. F. (1995). Effectiveness of self-management on attentional behavior and reading comprehension for children with attention deficit disorder. *Child & Family Behavior Therapy*, *17*(2), 1-17.

Fantuzzo, J., Harrell, K., & McLeod, M. (1980). Across-subject generalization of attending behavior as a function of self-regulation training. *Child Behavior Therapy*, *1*(4), 313-321.

Firman, K. B., Beare, P., & Loyd, R. (2002). Enhancing Self-Management in Students with Mental Retardation: Extrinsic versus Intrinsic Procedures. *Education and Training in Mental Retardation and Developmental Disabilities*,*37*(2), 163-171.

Gilberts, G. H., Agran, M., Hughes, C., & Wehmeyer, M. (2001). The effects of peer delivered self-monitoring strategies on the participation of students with severe disabilities in general education classrooms. *Research and Practice for Persons with Severe Disabilities*, *26*(1), 25-36.

Glynn, E. L., & Thomas, J. D. (1974). Effect of cueing on self-control of classroom behavior. *Journal of Applied Behavior Analysis*, *7*(2), 299-306.

Gumpel, T. P., & David, S. (2000). Exploring the Efficacy of Self-Regulatory Training as a Possible Alternative to Social Skills Training. *Behavioral Disorders*, *25*(2), 131-141.

Gureasko-Moore, S., DuPaul, G., & White, G. P. (2007). Self-Management of Classroom Preparedness and Homework: Effects on School Functioning of Adolescent with Attention Deficit Hyperactivity Disorder. *School Psychology Review*, *36*(4), 647-664.

Hallahan, D. P., Lloyd, J. W., Kneedler, R. D., & Marshall, K. J. (1982). A comparison of the effects of self-versus teacher-assessment of on-task behavior. *Behavior Therapy*, *13*(5), 715-723.

Hallahan, D. P., Lloyd, J., Kosiewicz, M. M., Kauffman, J. M., & Graves, A. W. (1979). Self-monitoring of attention as a treatment for a learning disabled boy's off-task behavior. *Learning Disability Quarterly*, 24-32.

Hallahan, D. P., Marshall, K. J., & Lloyd, J. W. (1981). Self-recording during group instruction: Effects on attention to task. *Learning Disability Quarterly*, 407-413.

Harris, K. R. (1986). Self-monitoring of attentional behavior versus self-monitoring of productivity: effects on on-task behavior and academic response rate among learning disabled children. *Journal of Applied Behavior Analysis*,*19*(4), 417-423.

Harris, K. R., Friedlander, B. D., Saddler, B., Frizzelle, R., & Graham, S. (2005). Self-Monitoring of Attention Versus Self-Monitoring of Academic Performance Effects

Among Students with ADHD in the General Education Classroom. *The Journal of Special Education*, *39*(3), 145-157.

Harris, K. R., Graham, S., Reid, R., McElroy, K., & Hamby, R. S. (1994). Self-monitoring of attention versus self-monitoring of performance: Replication and cross-task comparison studies. *Learning Disability Quarterly*, 121-139.

Hogan, S., & Prater, M. A. (1993). The effects of peer tutoring and self-management training on on-task, academic, and disruptive behaviors. *Behavioral Disorders*. 118-128.

Horner, R. H., & Brigham, T. A. (1979). The Effects of Self-Management Procedures on the Study Behavior of Two Retarded Children. *Education and Training of the Mentally Retarded*, *14*(1), 18-24.

Houghton, S. (1989). Improving social behaviour and academic performance of a secondary school pupil through self-recording: a replication of Merrett and Blundell. *Educational Psychology*, *9*(3), 239-245.

Hughes, C. A., & Boyle, J. R. (1991). Effects of Self-Monitoring for On-Task Behavior and Task Productivity on Elementary Students with Moderate Mental Retardation. *Education and Treatment of Children*, *14*(2), 96-111.

Hughes, C., & Scott, S. V. (1997). Teaching self-management in employment settings. *Journal of Vocational Rehabilitation*, *8*(1), 43-53.

Hughes, C., Copeland, S. R., Agran, M., Wehmeyer, M. L., Rodi, M. S., & Presley, J. A. (2002). Using self-monitoring to improve performance in general education high school classes. *Education and Training in Mental Retardation and Developmental Disabilities*, *37*(3), 262-272.

Hutchinson, S. W., Murdock, J. Y., Williamson, R. D., & Cronin, M. E. (2000). Self-Recording PLUS Encouragement Equals Improved Behavior. *Teaching Exceptional Children*, *32*(5), 54-58.

Koegel, L. K., Harrower, J. K., & Koegel, R. L. (1999). Support for children with developmental disabilities in full inclusion classrooms through self-management. *Journal of Positive Behavior Interventions*, *1*(1), 26-34.

Koegel, R. L., & Frea, W. D. (1993). Treatment of social behavior in autism through the modification of pivotal social skills. *Journal of Applied Behavior Analysis*, *26*(3), 369-377.

Lee, C., & Tindal, G. A. (1994). Self-recording and goal-setting: Effects on on-task and math productivity of low-achieving Korean elementary school students. *Journal of Behavioral Education*, *4*(4), 459-479.

Levendoski, L. S., & Cartledge, G. (2000). Self-monitoring for elementary school children with serious emotional disturbances: Classroom applications for increased academic responding. *Behavioral Disorders*, *25*(3), 211-24.

Lloyd, J. W., Hallahan, D. P., Kosiewicz, M. M., & Kneedler, R. D. (1982). Reactive effects of self-assessment and self-recording on attention to task and academic productivity. *Learning Disability Quarterly*, 216-227.

Lloyd, M. E., & Hilliard, A. M. (1989). Accuracy of self-recording as a function of repeated experience with different self-control contingencies. *Child & family behavior therapy*, *11*(2), 1-14.

Maag, J. W., Reid, R., & DiGangi, S. A. (1993). Differential effects of self –monitoring attention, accuracy, and productivity. *Journal of Applied Behavior Analysis*, *26*(3), 329-344.

Maag, J. W., Rutherford, R. B., & Digangi, S. A. (1992). Effects of self-monitoring and contingent reinforcement on on-task behavior and academic productivity of learning-disabled students: A social validation study. *Psychology in the Schools*, *29*(2), 157-172.

Mancina, C., Tankersley, M., Kamps, D., Kravits, T., & Parrett, J. (2000). Brief Reports Brief Report: Reduction of Inappropriate Vocalizations for a Child with Autism Using a Self-Management Treatment Program. *Journal of Autism and Developmental Disorders*, *30*(6), 599-606.

Marshall, K. J., Lloyd, J. W., & Hallahan, D. P. (1993). Effects of training to increase self-monitoring accuracy. *Journal of Behavioral Education*, *3*(4), 445-459.

Martella, R. C., Leonard, I. J., Marchand-Martella, N. E., & Agran, M. (1993). Self-monitoring negative statements. *Journal of Behavioral Education*, *3*(1), 77-86

Mathes, M. Y., & Bender, W. N. (1997). The effects of self-monitoring on children with attention-deficit/hyperactivity disorder who are receiving pharmacological interventions. *Remedial and Special Education*, *18*(2), 121-128.

McCarl, J. J., Svobodny, L., & Beare, P. L. (1991). Self-recording in a classroom for students with mild to moderate mental handicaps: Effects on productivity and on-task behavior. *Education & Training in Mental Retardation, 26(1), 79-88.*

McDougall, D., & Brady, M. P. (1995). Using audio-cued self-monitoring for students with severe behavior disorders. *The Journal of Educational Research*,*88*(5), 309-317.

McGoey, K. E., Prodan, T., & Condit, N. (2007). Examining the effects of teacher and self-evaluation of disruptive behavior via school-home notes for two young children in kindergarten. *Journal of Early and Intensive Behavior Intervention*, *3*(4), 365-376.

Minner, S. (1990). Use of a self-recording procedure to decrease the time taken by behaviorally disordered students to walk to special classes. *Behavioral Disorders*. 15(4), 210-216.

Moore, D. W., Prebble, S., Robertson, J., Waetford, R., & Anderson, A. (2001). Self-recording With Goal Setting: a self-management programme for the classroom. *Educational Psychology*, *21*(3), 255-265.

Morrison, L., Kamps, D., Garcia, J., & Parker, D. (2001). Peer mediation and monitoring strategies to improve initiations and social skills for students with autism. *Journal of Positive Behavior Interventions*, *3*(4), 237-250.

Nelson, J. R., Smith, D. J., & Colvin, G. (1995). The effects of a peer-mediated self-evaluation procedure on the recess behavior of students with behavior problems. *Remedial and Special Education*, *16*(2), 117-126.

Newman, B., Buffington, D. M., Hemmes, N. S. (1996). Self-Reinforcement Used To Increase the Appropriate Conversation of Autistic Teenagers. *Education and Training in Mental Retardation and Developmental Disabilities*, *31*(4), 304-09.

Ninness, H. A. C, Fuerst, J., Rutherford, RD, & Glenn, S. S. (1991). Effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. *Journal of Applied Behavior Analysis*, *24*(3), 499-508.

Ninness, H. C., Ellis, J., Miller, W. B., Baker, D., & Rutherford, R. (1995). The effect of a self-management training package on the transfer of aggression control procedures in the absence of supervision. *Behavior Modification*, *19*(4), 464-490.

Noland, S. A., Arnold, J., & Clement, P. W. (1980). Self-reinforcement by under-achieving, under-controlled girls. *Psychological reports*, *47*(2), 671-678.

O'Reilly, M., Tiernan, R., Lacey, C., Hillery, J., & Gardiner, M. (2002). Use of Self-Monitoring and Delayed Feedback To Increase On-Task Behavior in a Post-Institutionalized Child within Regular Classroom Settings. *Education and Treatment of Children*, *25*(1), 91-102.

Parker, D., & Kamps, D. (2011). Effects of task analysis and self-monitoring for children with autism in multiple social settings. *Focus on Autism and Other Developmental Disabilities*, *26*(3), 131-142.

Peterson, L. D., Young, K. R., Salzberg, C. L., West, R. P., & Hill, M. (2006). Using Self-Management Procedures to Improve Classroom Social Skills in Multiple General Education Settings. *Education and treatment of children,29*(1), 1-21.

Prater, M. A., Hogan, S., & Miller, S. R. (1992). Using self-monitoring to improve on-task behavior and academic skills of an adolescent with mild handicaps across special and regular education settings. *Education & Treatment of Children*, 15(1), 43-55.

Prater, M. A., Joy, R., Chilman, B., Temple, J., & Miller, S. R. (1991). Self-monitoring of on-task behavior by adolescents with learning disabilities.*Learning Disability Quarterly*, 164-177.

Quillivan, C. C. , Skinner, C. H. , Embree, M. H. , Whithed, D. and Ballard, D. 2011. Using a cell phone to prompt a kindergarten student to self-monitor off-task/disruptive behavior. Journal of Evidence-Based Practices for Schools, 12: 131–146.

Rafferty, L. A., & Raimondi, S. L. (2009). Self-monitoring of attention versus self-monitoring of performance: Examining the differential effects among students with emotional disturbance engaged in independent math practice.*Journal of Behavioral Education*, *18*(4), 279-299.

Reinecke, D. R., Newman, B., & Meinberg, D. L. (1999). Self-Management of Sharing in Three Pre-Schoolers with Autism. *Education and Training in Mental Retardation and Developmental Disabilities*, *34*(3), 312-17.

Roberts, R. N., & Nelson, R. O. (1981). The effects of self-monitoring on children's classroom behavior. *Child Behavior Therapy, 3*(2/3), 105-120.

Rock, M. L. (2005). Use of strategic self-monitoring to enhance academic engagement, productivity, and accuracy of students with and without exceptionalities. *Journal of Positive Behavior Interventions*, *7*(1), 3-17.

Rock, M. L., & Thead, B. K. (2007). The effects of fading a strategic self-monitoring intervention on students’ academic engagement, accuracy, and productivity. *Journal of Behavioral Education*, *16*(4), 389-412.

Rooney, K. J., & Hallahan, D. P. (1988). The Effects of Self-Monitoring on Adult Behavior and Student Independence. *Learning Disabilities Research*, *3*(2), 88-93.

Rooney, K. J., Hallahan, D. P., & Lloyd, J. W. (1984). Self-recording of attention by learning disabled students in the regular classroom. *Journal of Learning Disabilities*, *17*(6), 360-364.

Rumsey, I., & Ballard, K. D. (1985). Teaching self-management strategies for independent story writing to children with classroom behaviour difficulties.*Educational Psychology*, *5*(2), 147-157.

Rusch, F. R. (1988). Teaching a Student with Severe Handicaps to Self-Instruct: A Brief Report. *Education and Training in Mental Retardation*, *23*(1), 51-58.

Sainato, D. M., Goldstein, H., & Strain, P. S. (1992). Effects of self-evaluation on preschool children's use of social interaction strategies with their classmates with autism. *Journal of Applied Behavior Analysis*, *25*(1), 127-141.

Sainato, D. M., Strain, P. S., Lefebvre, D., & Rapp, N. (1990). Effects of self-evaluation on the independent work skills of preschool children with disabilities. *Exceptional Children*. 56(6), 540-549

Shearer, D. D., Kohler, F. W., Buchan, K. A., & McCullough, K. M. (1996). Promoting independent interactions between preschoolers with autism and their nondisabled peers: An analysis of self-monitoring. *Early Education and Development*, *7*(3), 205-220.

Smith, B. W., & Sugai, G. (2000). A self-management functional assessment-based behavior support plan for a middle school student with EBD. *Journal of Positive Behavior Interventions*, *2*(4), 208-217.

Smith, D. J. Nelson, J. R., Young, K. R., West, R. P. (1992). The effect of a self-management procedure on the classroom and academic behavior of students with mild handicaps. *School Psychology Review*, *21*(1), 59-72.

Smith, D. J., Young, K. R., West, R. P., & Morgan, D. P. (1988). Reducing the disruptive behavior of junior high school students: A classroom self-management procedure. *Behavioral Disorders*. 13(4), 231-239.

Stevenson, H. C., & Fantuzzo, J. W. (1984). Application of the “Generalization Map” to a self-control intervention with school aged children. *Journal of applied behavior analysis*, *17*(2), 203-212.

Sugai, G., & Rowe, P.(1984). The effect of self-recording on out-of-seat behavior of an EMR student. *Education and Training of the Mentally Retarded,19,* 23-28.

Terenzi, C. M., Ervin, R. A., & Hoff, K. E. (2010). Classwide self-management of rule following: Effect on the on-task and disruptive behaviors of three students with specific learning disabilities and attention-deficit/hyperactivity disorder. *Journal of Evidence-Based Practices for Schools*. 11(2), 87-116.

Todd, A. W., Horner, R. H., & Sugai, G. (1999). Self-Monitoring and Self-Recruited Praise Effects on Problem Behavior, Academic Engagement, and Work Completion in a Typical Classroom. *Journal of Positive Behavior Interventions*, *1*(2), 66-122.

Waller, R. J., Albertini, C. L., & Waller, K. S. (2011). Self-Monitoring of Performance to Promote Accurate Work Completion: A Functional Based Intervention for a 4th Grade Student Presenting Challenging Behavior. *Advances in School Mental Health Promotion*, *4*(1), 52-60.

Wolfe, L. H., Heron, T. E., & Goddard, Y. L. (2000). Effects of self-monitoring on the on-task behavior and written language performance of elementary students with learning disabilities. *Journal of Behavioral Education*, *10*(1), 49-73.

Wood, S. J., Murdock, J. Y., Cronin, M. E., Dawson, N. M., & Kirby, P. C. (1998). Effects of self-monitoring on on-task behaviors of at-risk middle school students. *Journal of Behavioral Education*, *8*(2), 263-279.